Performability analysis of the second order semi-Markov chains in state and duration for wind speed modeling

Guglielmo D’Amico\textsuperscript{a}, Filippo Petroni\textsuperscript{b}, Flavio Prattico\textsuperscript{c}

\textsuperscript{a}Dipartimento di Scienze del Farmaco, Università G. d’Annunzio, 66013 Chieti, Italy
E-mail: g.damico@unich.it

\textsuperscript{b}Dipartimento di Scienze Economiche e Aziendali, Università degli studi di Cagliari,
09123 Cagliari, Italy
E-mail: fpetroni@unica.it

\textsuperscript{c}Dipartimento di Ingegneria Meccanica, Energetica e Gestionale, Università degli studi
dell’Aquila, 67100 L’Aquila, Italy
E-mail: flavioprattico@gmail.com

Abstract

In this paper second order semi-Markov reward models are presented and equations for the higher order moments of the reward process are presented for the first time and applied to wind energy production. A real application is executed by considering a database, freely available from the web, in which are included wind speed data taken from L.S.I. - Lastem station (Italy) and sampled every 10 minutes. We compute the expected total energy produced by using the blade Aircon HAWT - 10 kW.

Keywords: semi-Markov chains, synthetic time series, autocorrelation

1. Introduction

Nowadays the subject of wind speed modeling is becoming increasingly important. Wind power generation companies are intensely interested in quantifying the energy which can be extracted from wind by using different types of wind turbines in a given location.

Models to predict future wind speed and power production have been proposed extensively in last two decades. In general the wind speed is divided into a finite number of states. Transitions between states occur randomly in time. The installed blade produces energy depending on its technical characteristic and on the wind speed status.

Markov chains have been extensively used to model the behavior of wind
speed data. Several authors have discussed the use of techniques of Markov processes in wind speed modeling, see e.g., Shamshad et al. (2005), Nfaoui et al. (2004) and Youcef et al. (2003). The Markovian assumption has, especially in the modeling of wind speed, several flaws. In discrete time, waiting times in a state before making a transition to another state are geometrically distributed and consequently the memoryless property applies. This leads to a great simplification of the model which is unable to reproduce correctly the statistical properties of the real wind speed process.

Semi-Markov chains do not have this constraint, because the waiting time distribution function in the states can be of any type and this allow the data to speak for themselves without any restriction. D’Amico et al. (2011) was the first paper were semi-Markov chains were applied in the modeling of wind speed. In that paper first and second order semi-Markov models were proposed with the aim of generate reliable synthetic wind speed data. It was shown that all the semi-Markov models perform better than the Markov chain model in reproducing the statistical properties of wind speed data. In particular, the model recognized as being the more suitable is the second order semi-Markov model in state and duration.

One of the purposes of this paper is to provides methods for computing the accumulated energy produced by a blade in a temporal interval $[0,T]$. To this end we introduce semi-Markov reward processes.

Semi-Markov reward processes were applied in several domains, for example De Dominicis and Manca (1986) applied non-homogeneous semi-Markov reward processes to insurance disability problems. In Stenberg et al. (2007) backward semi-Markov reward processes were considered for calculating any integer moment of the reward process.

In this paper we give a generalization of these results by defining the second order semi-Markov reward process in state and duration and giving relations for computing the higher order moments of this process. We apply the theoretical results in computing the accumulated energy produced by a blade during a bounded time interval. The expected total energy produced gives important information on the feasibility of the investment in a wind farm and the riskiness of the investment can be measured in terms of variance, skewness, and kurtosis of the reward process. Finally notice that the technological characteristics of different blades are captured by the permanence reward and consequently we are able to choose among different blades to be installed at a given location. Additionally, we propose to employ a matrix notation that makes calculations easier and also provides a compact
form for equations of moments of the reward process.

2. The second order semi-Markov chain in state and duration

In this section we describe briefly the second order semi-Markov chain in state and duration, see D’Amico et al. (2012) for additional results.

Let consider a finite set of states \( E = \{1, 2, ..., S\} \) in which the system can be into and a complete probability space \((\Omega, F, \mathbb{P})\) on which we define the following random variables:

\[ J_n : \Omega \to E, \quad T_n : \Omega \to \mathbb{N}. \tag{1} \]

They denote the state occupied at the \( n \)-th transition and the time of the \( n \)-th transition, respectively. To be more concrete, by \( J_n \) we denote the wind speed at the \( n \)-th transition and by \( T_n \) the time of the \( n \)-th transition of the wind speed.

We assume that

\[
\mathbb{P}[J_{n+1} = j, T_{n+1} - T_n = t | \sigma(J_s, T_s), J_n = k, J_{n-1} = i, T_n - T_{n-1} = x, 0 \leq s \leq n] = \mathbb{P}[J_{n+1} = j, T_{n+1} - T_n = t | J_n = k, J_{n-1} = i, T_n - T_{n-1} = x] := x_{q_{i,k,j}}(t). \tag{2}
\]

Relation (2) asserts that, the knowledge of the values \( J_n, J_{n-1}, T_n - T_{n-1} \) suffices to give the conditional distribution of the couple \( J_{n+1}, T_{n+1} - T_n \) whatever the values of the past variables might be. Therefore to make probabilistic forecasting we need the knowledge of the last two visited state and the duration time of the transition between them. For this reason we called this model a second order semi-Markov chains in state and duration.

It should be remarked that in the paper by Limnios and Oprešan (2003) were defined nth order semi-Markov chains in continuous time. Anyway the dependence was only on past states and not on durations.

The conditional probabilities

\[ x_{q_{i,k,j}}(t) = \mathbb{P}[J_{n+1} = j, T_{n+1} - T_n = t | J_n = k, J_{n-1} = i, T_n - T_{n-1} = x] \]

are stored in a matrix of functions \( q = (x_{q_{i,k,j}}(t)) \) named the second order kernel (in state and duration). The element \( x_{q_{i,k,j}}(t) \) represents the probability that next wind speed will be in speed \( j \) at time \( t \) given that the current wind speed is \( k \) and the previous wind speed state was \( i \) and the duration in wind speed \( i \) before of reaching wind speed \( k \) was equal to \( x \) units of time.
From the knowledge of the kernel we can define the cumulated second order kernel probabilities:

$$xQ_{i,k,j}(t) := P[J_{n+1} = j, T_{n+1} - T_n \leq t | J_n = k, J_{n-1} = i, T_n - T_{n-1} = x]$$

$$= \sum_{s=1}^{t} xq_{i,k,j}(s).$$

(3)

The process \(\{J_n\}\) is a second order Markov chain with state space \(E\) and transition probability matrix \(xP = xQ(\infty)\). We shall refer to it as the embedded Markov chain.

Define the unconditional waiting time distribution function in states \(k\) coming from state \(i\) with duration \(x\) as

$$xH_{i,k}(t) := P[T_{n+1} - T_n \leq t | J_n = k, J_{n-1} = i, T_n - T_{n-1} = x] = \sum_{j \in E} xQ_{i,k,j}(t).$$

(4)

The conditional cumulative distribution functions of the waiting time in each state, given the state subsequently occupied is defined as

$$xG_{i,k,j}(t) = P[T_{n+1} - T_n \leq t | J_n = k, J_{n-1} = i, J_{n+1} = j, T_n - T_{n-1} = x]$$

$$= \frac{1}{xP_{i,k,j}} \sum_{s=1}^{t} xq_{i,k,j}(s) \cdot 1_{\{xP_{i,k,j} \neq 0\}} + 1_{\{xP_{i,k,j} = 0\}}$$

(5)

Define by \(N(t) = \sup\{n : T_n \leq t\} \ \forall t \in \mathbb{N}\). We define the second order (in state and duration) semi-Markov chain as \(Z(t) = (Z^1(t), Z^2(t)) = (J_{N(t)-1}, J_{N(t)})\).

For this model ordinary transition probability functions and transition probabilities with initial and final backward recurrence times were defined and computed D’Amico et al. (2012).

3. The Reward Model

In this section by following the line of research in Stenberg et al. (2007) we determine recursive equations for higher order moments of the second order semi-Markov reward chain in state and duration.

Let \(\xi(t)\) denote the accumulated discounted reward during the time interval \((0, t]\) defined by the following relation,

$$\xi(t) = \sum_{0 < u \leq t} x_{N(u)-2}^{\psi_1J_{N(u)-2}, J_{N(u)-1}; J_{N(u)}(B(u))} e^{-\delta u}$$

(6)
where: \( B(u) = u - T_{N(u)} \) is the backward recurrence time process, \( X_{N(u)} := T_{N(u)+1} - T_{N(u)} \) is the sojourn time in state \( J_{N(u)} \) before the \( N(u) + 1 \) transition and \( e^{-\delta} \) with \( \delta \in [0, 1] \) is a one period deterministic discount factor.

The reward \( x_{N(u)} \cdot J_{N(u)} \cdot J_{N(u)} \cdot B(u) \) is more general than those considered in Stenberg et al. (2007) because at the same time it is state dependent on the current state \( J_{N(u)} \) of the system; it depends on the last two visited states \( J_{N(u)-2}, J_{N(u)-1} \); it is duration dependent in the current state because it is a function of the backward process \( B(u) \) and finally it depends on the sojourn time in the past states being dependent on \( X_{N(u)} \).

Let also denote by \( x,v \xi_{i,k}(t) \) the random variable, which has the distribution the same with the conditional distribution for the random variable \( \xi(t) \) given that

\[
J_{N(0)-1} = k, J_{N(0)-2} = i, B(T_{N(0)-1}) = v, X_{N(0)-2} = x
\]

and let denote by \( x,v \Psi_{i,k}(t) := E[(x,v \xi_{i,k}(t))^n] \).

In order to propose a matrix notation that simplifies calculations and provides a compact form for next equations we need to introduce the adopted notation and products.

Given two \( m \times n \) matrices \( A \) and \( B \), their Hadamard matrix product \( \circ \) gives the \( m \times n \) matrix \( C \) whose generic element is given by:

\[
c_{ij} = a_{ij}b_{ij}.
\]

Let \( A \) be a \( m^2 \times m \) matrix and \( B \) be a \( m^2 \) column vector, their \( \otimes \) matrix product gives the \( m^2 \) column vector whose elements, for all \( i,k \in \{1,2,...,m\} \) are expressed by

\[
C_{(i-1):i:m+k} = \sum_{j=1}^{m} A_{(i-1):i:m+k,j} B_{(k-1):k:m+j,1}.
\]

The first order moment of the reward process \( x,v \xi_{i,k}(t) \) is computed in the following Theorem.

**Theorem 1.** The first order moment of the second order semi-Markov chain in state and duration satisfies the following matrix equation:

\[
x,v V^{(1)}(t) = x,v D(t) \otimes x,v \Phi(t) + \sum_{s=1}^{t} (x,v B(s) \cdot 1_{|E|}) \otimes x,v \Phi(s)
\]

\[
+ \sum_{s=1}^{t} e^{-\delta s} x,v B(s) \otimes x,s \otimes V^{(1)}(t - s)
\]

5
where \( \forall i, k \in E \)

\[
x_{i,v}V^{(1)}(t) = \left( x_{i,v}V_{(i-1)\cdot|E|+k}^{(1)}(t) \right) = \left( x_{i,v}V_{i,k}^{(1)}(t) \right),
\]

\[
x_{i,v}\Psi(v+u) = (x_{i,v}\psi_{(i-1)\cdot|E|+k}(v+u)e^{-\delta u}) = (x_{i,k;k}(v+u)e^{-\delta u}),
\]

\[
x_{i,v}\Psi(v+u) \text{ is } |E|^2 \times 1 \text{ and } x_{i,v}\tilde{\Psi}(t) = \sum_{u=1}^t x_{i,v}\Psi(v+u) \text{ is } |E|^2 \times 1,
\]

\[
x_{i,v}D(t) = (x_{i,v}D_{(i-1)\cdot|E|+k}(t)) = (x_{i,v}D_{i,k}(t)) = \left( \frac{1 - xH_{i,k}(t+v)}{1 - xH_{i,k}(v)} \right)
\]

\[
x_{i,v}B(s) = (x_{i,v}B_{(i-1)\cdot|E|+k,i}(s)) = \left( \frac{xq_{i,k,j}(s+v)}{1 - xH_{i,k}(v)} \right)
\]

and

\[
1_{[E]}^T = [1, 1, \ldots, 1]^T.
\]

**Proof:** Let consider the random variable \( x_{i,v}\xi_{i,k}(t) \). The time of next transition \( T_{N(0)+1} \) can be greater of \( t \) or not. Consequently it results that:

\[
x_{i,v}V_{i,k}^{(1)}(t) := E[x_{i,v}\xi_{i,k}(t)] = E[x_{i,v}\xi_{i,k}(t)1_{\{T_{N(0)+1}>t\}}] + E[x_{i,v}\xi_{i,k}(t)1_{\{T_{N(0)+1}\leq t\}}].
\]

In the case \( T_{N(0)+1} > t \) we have that

\[
x_{i,v}\xi_{i,k}(t) = \sum_{u=1}^t x_{i,v}\psi_{i,k;k}(u+v)e^{-\delta u}
\]

and this event occurs with probability

\[
\mathbb{P}(T_{N(0)+1} > t|T_{N(0)+1} > 0, T_{N(0)} = -v, J_{N(0)} = k, T_{N(0)-1} = -v-x, J_{N(0)-1} = i) = \frac{\mathbb{P}(T_{N(0)+1} > t, T_{N(0)+1} > 0, T_{N(0)} = -v|J_{N(0)} = k, T_{N(0)-1} = i, X_{N(0)-1} = x)}{\mathbb{P}(T_{N(0)+1} > 0, T_{N(0)} = -v|J_{N(0)} = k, T_{N(0)-1} = i, X_{N(0)-1} = x)} = \mathbb{P}(X_{N(0)} > t + v|J_{N(0)} = k, T_{N(0)-1} = i, X_{N(0)-1} = x) = \frac{1 - xH_{i,k}(t+v)}{1 - xH_{i,k}(v)} = x_{i,v}D_{i,k}(t).
\]

(10)
Then it results that

$$\mathbb{E}[x,v \xi_{i,k}(t)1_{(T_{N(0)+1} > t)}] = x,v \sum_{u=1}^{t} x\psi_{i,k}(u+v)e^{-\delta u}. \quad (11)$$

The right hand side of (11) can be expressed in matrix form as follows:

$$x,v \mathbf{D}(t) \otimes x,v \mathbf{\Psi}(t). \quad (12)$$

In the second case, when $T_{N(0)+1} \leq t$, if we consider the next visited state $J_{N(0)+1}$ and the time of next transition $T_{N(0)+1}$ we have:

$$x,v \xi_{i,k}(t) = \sum_{s=1}^{T_{N(0)+1}} x\psi_{i,k}(v + s')e^{-\delta s'} + v + T_{N(0)+1} \delta \xi_{i,k,J_{N(0)+1}}(t - s)e^{-\delta T_{N(0)+1}}. \quad (13)$$

The event $\{J_{N(0)+1} = j, T_{N(0)+1} = s\}$ occurs with probability

$$\mathbb{P}(J_{N(0)+1} = j, T_{N(0)+1} = s | T_{N(0)+1} > 0, T_{N(0)} = -v, J_{N(0)} = k, T_{N(0)-1} = -v - x, J_{N(0)-1} = i)$$

$$= \frac{\mathbb{P}(J_{N(0)+1} = j, T_{N(0)+1} = s, T_{N(0)+1} > 0, T_{N(0)} = -v | J_{N(0)} = k, T_{N(0)-1} = i, X_{N(0)-1} = x)}{\mathbb{P}(T_{N(0)+1} > 0, T_{N(0)} = -v | J_{N(0)} = k, T_{N(0)-1} = i, X_{N(0)-1} = x)}$$

$$= \frac{\mathbb{P}(X_{N(0)} > v | J_{N(0)} = k, T_{N(0)-1} = i, X_{N(0)-1} = x)}{\mathbb{P}(X_{N(0)} = v | J_{N(0)} = k, T_{N(0)-1} = i, X_{N(0)-1} = x)}$$

$$= xq_{i,k}^{j}(s+v) \frac{1}{1 - xH_{i,k}(v)} = x,vB(i-1,|E|+k,j)(s). \quad (14)$$

Notice that the random variable $v + T_{N(0)+1} \delta \xi_{i,k,J_{N(0)+1}}(t - s)$ is independent of the distribution of the joint random variable $(J_{N(0)+1}, T_{N(0)+1})$ because the accumulation process has the Markov property at transition times and consequently once the state $J_{N(0)+1}$ and the $T_{N(0)+1}$ are known its behaviour doesn’t depends on the distribution of $(J_{N(0)+1}, T_{N(0)+1})$. Then by taking the expectation in (13) we get

$$\mathbb{E}[x,v \xi_{i,k}(t) 1_{(T_{N(0)+1} \leq t)}]$$

$$= \sum_{j \in E} \sum_{s=1}^{t} xq_{i,k}^{j}(s+v) \frac{1}{1 - xH_{i,k}(v)} \sum_{s'=1}^{s} x\psi_{i,k}(v + s')e^{-\delta s'} \quad (15)$$

$$+ \sum_{j \in E} \sum_{s=1}^{t} xq_{i,k}^{j}(s+v) \frac{1}{1 - xH_{i,k}(v)} \times v + s,0V_{k,j}^{(1)}(t - s)e^{-\delta s}.$$
The right hand side of (15) can be expressed in matrix form as follows:

\[
x,v \mathbf{V}^{(1)}(t) = \sum_{s=1}^{t} (x,v \mathbf{B}(s) \cdot \mathbf{1}_{|E|}) \otimes x,v \tilde{\Psi}(s) \\
+ \sum_{s=1}^{t} e^{-\delta s} x,v \mathbf{B}(s) \otimes x,v \mathbf{V}^{(1)}(t-s).
\]  

(16)

A substitution of (12) and (16) in (8) concludes the proof.

By using similar techniques it is possible to get recursive equations for the higher order moments of the reward process.

First note that

\[
x,v \mathbf{V}^{(n)}_{i,k}(t) := \mathbb{E}[(x,v \xi_{i,k}(t))^n]
\]

\[= \mathbb{E}[(x,v \xi_{i,k}(t))^n 1_{T_N(0)+1 > t}] + \mathbb{E}[(x,v \xi_{i,k}(t))^n 1_{T_N(0)+1 \leq t}].
\]  

(17)

In the case \(T_N(0)+1 > t\) we have that

\[(x,v \xi_{i,k}(t))^n = \left( \sum_{u=1}^{t} x \psi_{i,k;}(u + v)e^{-\delta u} \right)^n
\]

(18)

and this event occurs with probability \(x,v D_{i,k}(t)\), see (10). Consequently it results that

\[\mathbb{E}[(x,v \xi_{i,k}(t))^n 1_{T_N(0)+1 > t}] = x,v D_{i,k}(t) \left( \sum_{u=1}^{t} x \psi_{i,k;}(u + v)e^{-\delta u} \right)^n.
\]  

(19)

In the opposite case, when \(T_N(0)+1 \leq t\), we have that

\[x,v \xi_{i,k}^{(n)}(t) = \left( \sum_{s'=1}^{T_N(0)+1} x \psi_{i,k;}(v + s')e^{-\delta s'} + v + T_N(0)+1 \xi_{k,J_N(0)+1} (t-s)e^{-\delta T_N(0)+1} \right)^n
\]

\[\left( \sum_{s'=1}^{T_N(0)+1} x \psi_{i,k;}(v + s')e^{-\delta s'} \right)^n + v + T_N(0)+1 \xi_{k,J_N(0)+1} (t-s)e^{-\delta T_N(0)+1}
\]

\[+ \sum_{l=1}^{n-1} \binom{n}{l} \left( \sum_{s'=1}^{s} x \psi_{i,k;}(v + s')e^{-\delta s'} \right)^{n-l} (v + T_N(0)+1) \xi_{k,J_N(0)+1} (t-s)e^{-\delta T_N(0)+1}
\]

(20)
The event \( \{ J_{N(0)+1} = j, T_{N(0)+1} = s \} \) occurs with probability \( x,v B(i-1) \cdot E_{j+k,j}(s) \). Then, by using the already mentioned independence between \( v+T_{N(0)+1}, d \xi_k, J_{N(0)+1}(t-s) \) and the joint random variable \( (J_{N(0)+1}, T_{N(0)+1}) \), by taking the expectation in (20) we get:

\[
\mathbb{E}[x,v \xi_{i,k}(t) \cdot 1_{\{T_{N(0)+1} \leq t\}}] = \\
+ \sum_{j \in E} \sum_{s=1}^{t} x q_{i,k,j}(s+v) \cdot \left( \sum_{s'-1}^{s} x \psi_{i,k,k}(v+s') e^{-\delta s'} \right)^n \\
+ \sum_{j \in E} \sum_{s=1}^{t} x q_{i,k,j}(s+v) \cdot \frac{1}{x H_{i,k}(v)} \cdot \left( \sum_{s'=1}^{s} x \psi_{i,k,k}(v+s') e^{-\delta s'} \right)^n \\
+ \sum_{j \in E} \sum_{s=1}^{t} x q_{i,k,j}(s+v) \cdot \frac{1}{x H_{i,k}(v)} \cdot \left( \sum_{s'=1}^{s} x \psi_{i,k,k}(v+s') e^{-\delta s'} \right)^n \\
\cdot \left( v+s,0 \right) \left( t-s \right) e^{-\delta s}.
\] (21)

If we substitute (19) and (21) in (17) and we represent the resulting expression in matrix form we obtain the following equation:

\[
x,v \Psi^{(n)}(t) = x,v D(t) \otimes x,v \tilde{\Psi}^{(n)}(t) + \sum_{s=1}^{t} (x,v B(s) \cdot 1_{|E|}) \otimes x,v \tilde{\Psi}^{(n)}(s) \\
+ \sum_{s=1}^{t} e^{-\delta s} x,v B(s) \otimes v+s,0 \Psi^{(n)}(t-s) \\
+ \sum_{s=1}^{t} \sum_{l=1}^{n-1} \left( \begin{array}{c} n \\ l \end{array} \right) x,v \Psi^{(n-l)}(s) \otimes (e^{-\delta s} x,v B(s) \otimes v+s,0 \Psi^{(l)}(t-s)) .
\] (22)

It should be remarked that if \( \forall i \in E \) and \( \forall x \in \mathbb{N} \) we have that

\[
 x q_{i,k,j}(t) = q_{k,j}(t), \quad x \psi_{i,k,j}(t) = \psi_{k,j}(t)
\]

then the second order semi-Markov reward chain model in state and duration collapses in a standard semi-Markov reward chain model and we recover exactly the results by Stenberg et al. (2007).

4. Application to real data

To check the validity of our model we perform a comparison of the behavior of real data and wind speeds generated through Monte Carlo simulations.
Based on the models described above. In this section we describe the database of real data used for the analysis, the method used to simulate synthetic wind speed time series and, at the end, we compare results from real and simulated data.

The data used in this analysis are freely available from http://www.lsi-lastem.it/meteo/page/dwnlddata.aspx. The station of L.S.I. -Lastem is situated in Italy at N 45 28’ 14,9” — E 9 22’ 19,9” and at 107 m of altitude. The station use a combined speed-direction anemometer at 22 m above the ground. It has a measurement range that goes from 0 to 60 m/s, a threshold of 0,38 m/s and a resolution of 0,05 m/s. The station processes the speed every 10 minute in a time interval ranging from 25/10/2006 to 28/06/2011. During the 10 minutes are performed 31 sampling which are then averaged in the time interval. In this work, we use the sampled data that represents the average of the modulus of the wind speed (m/s) without a considered specific direction. The database is then composed of about 230 thousands wind speed measures ranging from 0 to 16 m/s. The time series, together with its empirical probability density distribution are represented in Figure 1. The lower panel of the same figure shows that the wind speed follows a Weibul distribution.

Figure 1: Time series of wind speed and its empirical distribution.
To be able to model the wind speed as a semi-Markov chain the state space of wind speed has been discretized. In the example shown in this work we discretized wind speed into 8 states chosen to cover all the wind speed distribution. The state space is numerically represented by the set $E = \{0 − 1, 1 − 2, 2 − 3, 3 − 4, 4 − 5, 5 − 6, 6 − 7, > 7\} (m/s)$. From the discretized wind speeds we estimated the probabilities $P$ and $G$ to generate synthetic trajectories for a second order semi-Markov model in state and duration.

We give here the Monte Carlo algorithm used to simulate the trajectory in the time interval $[0, T]$ where $T$ is equal to the time length of the real data. The output of the algorithm consists in the successive visited states $\{J_0, J_1, \ldots\}$ and the jump times $\{T_0, T_1, \ldots\}$ up to the time $T$.

Denote by $X_n = T_{n+1} − T_n$.

1. Set $n = 0$, $J_{−1} = i$, $J_0 = k$, $X_{−1} = x$, horizon time $= T$;
2. Sample $J$ from $x_{n−1} p_{J_{n−1}, J_n}(\cdot)$ and set $J_{n+1} = J(\omega)$;
3. Sample $W$ from $x_{n−1} G_{J_{n−1}, J_n, J_{n+1}}(\cdot)$ and set $T_{n+1} = T_n + W(\omega)$;
4. If $T_{n+1} \geq T$ stop, else set $X_n = T_{n+1} − T_n$ and $n = n + 1$ and go to 2.

We show in Figure 2 a short sample of 160 hours of the real and simulated trajectories just for comparison reason. We apply our model to a real case of energy production. For this reason we choose a commercial wind turbine, a 10 kW Aricon HAWT with a power curve given in Figure 3. The power curve of a wind turbine represents how it produces energy as a function of the wind speed. In this case we have no production of energy in the interval 0-2 m/s, the wind turbine produces energy linearly from 3 m/s to 10 m/s, than, with increasing wind speed the production remain constant until the limit of wind speed in which the wind turbine is stopped for structural reason. Through this power curve we transform the wind speed of our real and synthetic data into the equivalent energy produced at each different state of wind speed, in order to validate our model in a real case of energy production. For this reason, to check the proper functioning of the proposed model, we do several comparisons between real and synthetic data. In Figure 4 is showed the daily energy production of real and synthetic data. Another comparison between real and synthetic data is made by the cumulated energy produced in a time interval. In Figure 5 are showed these values as a function of time. The
Figure 2: Comparison of the time series of wind speed from real and simulated data

Figure 3: Power curve of the 10 kW Aricon HAWT wind turbine.
Figure 4: Comparison of the daily energy production for the real and simulated data. The error bar represent one standard deviation.

Figure 5: Comparison between the cumulated energy produced by real and simulated data. The two dashed line represents the standard deviation of the synthetic data. It is interesting to note that the trend of the real cumulated energy produced remains almost inside the area formed by the two dashed line. We generate 500 different trajectories and for each one we calculate the cumulated energy for all the length of the time series. In Figure 6 is plotted the distribution of the cumulated energy produced by the different trajectories. The black point represents the energy produced by the real data. The figure shows that simulated and real data give the same value for produced energy in the time interval within statistical error.
5. Conclusion

The wind is a very unstable phenomenon characterized by a sequence of lulls and sustained speeds, and a good wind generator must be able to reproduce such sequences. We have then modeled wind speed through a second order semi-Markov model. Our work is motivated by the presence of persistence in the wind speed process. The purposes of this paper is to provide theoretical methods for computing the accumulated energy produced by a blade in a temporal interval $[0, T]$. We have shown, by means of Monte Carlo simulations, that the proposed model is able to reproduce well the behavior of real data as far as energy production from wind speed is concerned.

References


First and second order semi-Markov chains for wind speed modeling

Guglielmo D’Amico\textsuperscript{1}, Filippo Petroni\textsuperscript{2}, and Flavio Prattico\textsuperscript{3}

\textsuperscript{1} Dipartimento di Scienze del Farmaco, Università G. d’Annunzio, 66013 Chieti, Italy, (E-mail: g.damico@unich.it)
\textsuperscript{2} Dipartimento di Scienze Economiche e Aziendali, Università degli studi di Cagliari, 09123 Cagliari, Italy, (E-mail: fpetroni@unica.it)
\textsuperscript{3} Dipartimento di Ingegneria Meccanica, Energetica e Gestionale, Università degli studi dell’Aquila, 67100 L’Aquila, Italy, (E-mail: flavioprattico@gmail.com)

Abstract. The increasing interest in renewable energy, particularly in wind, has given rise to the necessity of accurate models for the generation of good synthetic wind speed data. Markov chains are often used with this purpose but better models are needed to reproduce the statistical properties of wind speed data. We downloaded a database, freely available from the web, in which are included wind speed data taken from L.S.I. -Lastem station (Italy) and sampled every 10 minutes. With the aim of reproducing the statistical properties of this data we propose the use of three semi-Markov models. We generate synthetic time series for wind speed by means of Monte Carlo simulations. The time lagged autocorrelation is then used to compare statistical properties of the proposed models with those of real data and also with a synthetic time series generated though a simple Markov chain.

Keywords: semi-Markov chains, synthetic time series, autocorrelation.

1 Introduction

The increasing interest in renewable energy leads scientific research to find a better way to recover most of the available energy. Particularly, the maximum energy recoverable from wind is equal to 59.3\% of that available (Betz law) at a specific pitch angle and when the ratio between the wind speed in output and in input is equal to 1/3. A powerful system control of the pitch angle allows the wind turbine to recover better the energy in transient regime. A good stochastic model for wind speed is then needed to help both the optimization of turbine design and to assist the system control to predict the value of the wind speed to positioning the blades quickly and correctly. The possibility to have synthetic data of wind speed is a powerful instrument to assist designer to verify the structures of the wind turbines or to estimate the energy recoverable from a specific site. To generate synthetic data, Markov chains of first or higher order are often used \cite{1-3} but the search for a better model is still open. Approaching this issue, we applied new models which are generalization of Markov models. More precisely we applied first and second order semi-Markov models to generate synthetic wind speed time series.
The models are used to generate synthetic time series for wind speed by means of Monte Carlo simulations and the time lagged autocorrelation is used to compare statistical properties of the proposed models with those of real data and also with a time series generated through a simple Markov chain.

2 Wind speed modeling with semi-Markov chains

Semi-Markov chains are a generalization of Markov chains allowing the times between transitions to occur at random times according to any kind of distribution functions which may depend on the current and the next visited state. As it is well known, Markov chains have sojourn times between transitions geometrically distributed, for this reason the memoryless property is preserved and no duration effect is observed. The more general semi-Markov environment allows the possibility to use also no memoryless distributions and then can reproduce a duration effect. The duration effect affirms that the time the system is in a state influences its transition probabilities. The states of the process in our data are represented by different speed data, then in this paper we detect the presence of a duration effect in wind speed modeling and forecasting.

Here below we propose a semi-Markov model of order two in state and duration and we compare its performance with the Markov chain models often used to describe wind speed, see [1–3] and with some particular cases of semi-Markov chain models.

Let consider a finite set of states \( E = \{1, 2, \ldots, S\} \) in which the system can be into and a complete probability space \((\Omega, \mathcal{F}, P)\) on which we define the following random variables:

\[
J_n : \Omega \to E, \quad T_n : \Omega \to \mathbb{N}.
\] (1)

They denote the state occupied at the n-th transition and the time of the n-th transition, respectively. To be more concrete, by \( J_n \) we denote the wind speed at the nth transition and by \( T_n \) the time of the nth transition in the wind speed. We do the following conditional independence assumption:

\[
P[J_{n+1} = j, T_{n+1} - T_n = t|\sigma(J_s, T_s), J_n = k, J_{n-1} = i, T_n - T_{n-1} = x, 0 \leq s \leq n]
= P[J_{n+1} = j, T_{n+1} - T_n = t|J_n = k, J_{n-1} = i, T_n - T_{n-1} = x] := xq_{i,k,j}(t).
\] (2)

Relation (2) asserts that, the knowledge of the values \( J_n, J_{n-1}, T_n - T_{n-1} \) suffices to give the conditional distribution of the couple \( J_{n+1}, T_{n+1} - T_n \) whatever the values of the past variables might be. Therefore to make probabilistic forecasting we need the knowledge of the last two visited state and the duration time of the transition between them. For this reason we called this model a second order semi-Markov chains in state and duration.

The conditional probabilities

\[
xq_{i,k,j}(t) = P[J_{n+1} = j, T_{n+1} - T_n = t|J_n = k, J_{n-1} = i, T_n - T_{n-1} = x]
\]
are stored in a matrix of functions \( q = (q_{i,k,j}(t)) \) called the second order kernel (in state and duration). The element \( q_{i,k,j}(t) \) represents the probability that next wind speed will be in speed \( j \) at time \( t \) given that the current wind speed is \( k \) and the previous wind speed state was \( i \) and the duration in wind speed \( i \) before of reaching wind speed \( k \) was equal to \( x \) units of time.

We can define the cumulated second order kernel:

\[
x Q_{i,k,j}(t) := \mathbb{P}[J_{n+1} = j, T_{n+1} - T_n \leq t | J_n = k, J_{n-1} = i, T_n - T_{n-1} = x]
\]

\[
= \sum_{s=1}^{t} x q_{i,k,j}(s).
\]  

(3)

The process \( \{J_n\} \) is a second order Markov chain with state space \( E \) and transition probability matrix \( P = x Q(\infty) \). We shall refer to it as the embedded Markov chain.

Define the unconditional waiting time distribution function in states \( i \) and \( k \) with duration \( x \) as

\[
x H_{i,k}(t) := \mathbb{P}[T_{n+1} - T_n \leq t | J_n = k, J_{n-1} = i, T_n - T_{n-1} = x] = \sum_{j \in E} x Q_{i,k,j}(t).
\]

(4)

The conditional cumulative distribution functions of the waiting time in each state, given the state subsequently occupied is defined as

\[
x G_{i,k,j}(t) = \mathbb{P}[T_{n+1} - T_n \leq t | J_n = k, J_{n-1} = i, J_{n+1} = j, T_n - T_{n-1} = x]
\]

\[
= \frac{1}{x p_{i,k,j}} \sum_{s=1}^{t} x q_{i,k,j}(s) \cdot 1_{\{x p_{i,k,j} \neq 0\}} + 1_{\{x p_{i,k,j} = 0\}}
\]  

(5)

Define by \( N(t) = \sup\{n : T_n \leq t\} \forall t \in \mathbb{N} \). We define the second order (in state and duration) semi-Markov chain as \( Z(t) = (Z^1(t), Z^2(t)) = (J_{N(t)-1}, J_{N(t)}) \).

If we define, \( \forall i, k, j \in E \), and \( t \in \mathbb{N} \), the semi-Markov transition probabilities:

\[
x \phi_{i,k,h,j}(t) := \mathbb{P}[J_{N(t)} = j, J_{N(t)-1} = h | J_0 = k, J_1 = i, T_0 = 0, T_0 - T_{-1} = x],
\]

then the following system of equations is verified:

\[
x \phi_{i,k,h,j}(t) = 1_{\{i=h, k=j\}} (1 - x H_{i,k}(t)) + \sum_{r \in E} \sum_{s=1}^{t} x q_{i,k,r}(s) \cdot x \phi_{k,r,h,j}(t-s).
\]  

(7)

The proof of equation (7) is not given here because it is a particular case of the equation established and proved later.

To detect the duration effects let us introduce the backward recurrence time process defined for each time \( t \in \mathbb{N} \) by:

\[
B(t) = t - T_{N(t)}.
\]

(8)
If the semi-Markov process $Z(t)$ indicates the wind speed at time $t$, the backward process $B(t)$ indicates the time since the last transition, that is from how long time the wind speed is at the value $Z(t)$.

To quantify the duration effect in our second order semi-Markov model, let us define the following probabilities:

\[
\phi_{i,k,h,j}(v;v',t) := P[J_N(t) = j, B(t) = v', J_{N(t)-1} = h | J_0 = k, J_{-1} = i, B(0) = v, T_0 - T_{-1} = x].
\]

Expression (10) gives the probability that the wind speed will enter in the state $j$ at time $t - v'$ coming from state $h$ and will remains inside the state $j$ without any other transition up to the time $t$ given that at the present the wind speed is $k$ and it entered into this state with the last transition $v$ periods before coming from a wind speed equal to $i$ with a duration in $i$ of $x$ periods.

**Proposition 1.** The relation (10) represents the evolution equation of (9):

\[
\phi_{i,k,h,j}(v;v',t) = 1_{\{i=h,k=j,v'=v+t\}} \frac{[1 - H_{i,k}(t + v)]}{[1 - H_{i,k}(v)]} + \sum_{r \in E} \sum_{s=1}^{t-v'} \frac{x q_{i,k,r}(s + v)}{[1 - H_{i,k}(v)]} \phi_{r,h,j}(v';v',t-s).
\]

**Proof:** We have

\[
\phi_{i,k,h,j}(v;v',t)
\]

\[
= P[J_N(t) = j, B(t) = v', J_{N(t)-1} = h, T_1 > t | J_0 = k, J_{-1} = i, B(0) = v, T_0 - T_{-1} = x]
\]

\[
+ P[J_N(t) = j, B(t) = v', J_{N(t)-1} = h, T_1 \leq t | J_0 = k, J_{-1} = i, B(0) = v, T_0 - T_{-1} = x].
\]

Observe that

\[
P[J_N(t) = j, B(t) = v', J_{N(t)-1} = h, T_1 > t | J_0 = k, J_{-1} = i, B(0) = v, T_0 - T_{-1} = x]
\]

\[
= P[k = j, v' = v + t, i = h | T_1 > t, J_0 = k, J_{-1} = i, B(0) = v, T_0 - T_{-1} = x]
\]

\[
\cdot P[T_1 > t | J_0 = k, J_{-1} = i, T_0 - T_{-1} = x, T_0 = -v, T_1 > 0]
\]

\[
= 1_{\{k = j, i = h, v' = v + t\}} P[T_1 > t | J_0 = k, J_{-1} = i, T_0 - T_{-1} = x, T_0 = -v] P[T_1 > 0 | J_0 = k, J_{-1} = i, T_0 - T_{-1} = x, T_0 = -v]
\]

\[
= 1_{\{k = j, i = h, v' = v + t\}} \frac{1 - x H_{i,k}(t + v)}{1 - x H_{i,k}(v)}.
\]
The second addend on the right hand side of (11) can be represented as follows:

\[
\sum_{r \in E} \sum_{s=1}^{t-v'} P[J_{N(t)} = j, B(t) = v', J_{N(t)} = h, J_1 = r, T_1 = s| J_0 = k, J_{-1} = i, B(0) = v, T_0 - T_{-1} = x] \cdot P[J_1 = r, T_1 = s| J_0 = k, J_{-1} = i, B(0) = v, T_0 - T_{-1} = x]
\]

\[
= \sum_{r \in E} \sum_{s=1}^{t-v'} P[J_{N(t)} = j, B(t) = v', J_{N(t)} = h| J_1 = r, J_0 = k, B(s) = 0, T_1 - T_0 = s + v] \cdot P[J_1 = r, T_1 - T_0 = s + v| J_0 = k, J_{-1} = i, T_0 = -v, T_1 > 0, T_0 - T_{-1} = x]
\]

\[
= \sum_{r \in E} \sum_{s=1}^{t-v'} b\phi_{k,r,h,j}^{b}(0; v', t - s) \cdot \sum_{s=1}^{t-v'} \frac{b\phi_{k,r,h,j}^{b}(0; v', t - s)}{1 - xH_0^{b}(v)}
\]

\[
\sum_{r \in E} \sum_{s=1}^{t-v'} \frac{b\phi_{i,k,h,j}^{b}(s + v)}{1 - xH_0^{b}(v)} \cdot \sum_{s=1}^{t-v'} b\phi_{k,r,h,j}^{b}(0; v', t - s).
\]

(14)

A substitution of (13) and (14) in (11) concludes the proof.

Obviously we have that

\[
\phi_{i,k,h,j}^{b}(v,t) := P[J_{N(t)} = j, J_{N(t)} = h| J_0 = k, J_{-1} = i, B(0) = v, T_0 = 0, T_0 - T_{-1} = x]
\]

\[
= \sum_{v' \geq 0} b\phi_{i,k,h,j}^{b}(v', t).
\]

(15)

It expresses the probability that the wind speed will be in the state \( j \) at time \( t \) coming from a wind speed equal to \( h \) given that at present the wind speed is \( k \) and it entered into this state with the last transition \( v \) periods before coming from a wind speed equal to \( i \) with a duration in \( i \) of \( x \) periods.

Moreover if \( v = 0 \) we obtain the equation (7).

It should be noted that our semi-Markov model of order two in state and duration contains as a special case the semi-Markov model of order two in state. The paper [4] proposed a \( n \)-order semi-Markov process (in state) in continuous time. The discrete time counterpart of order two (in state) is obtained through the following assumption:

\[
P[J_{n+1} = j, T_{n+1} - T_n = t| J_n = k, J_{n-1} = i, 0 \leq s \leq n]
\]

\[
= P[J_{n+1} = j, T_{n+1} - T_n = t| J_n = k, J_{n-1} = i] := q_{i,k,j}(t).
\]

Relation (16) asserts that, the knowledge of the values \( J_n, J_{n-1} \) suffices to give the conditional distribution of the couple \( J_{n+1}, T_{n+1} - T_n \) whatever
the values of the past variables might be. Therefore to make probabilistic forecasting we need the knowledge of the last two visited state. In the application we will refer to this model as the model named semi-Markov II.

Finally remark that if we assume that

\[ P[J_{n+1} = j, T_{n+1} - T_n = t | \sigma(J_s, T_s), J_n = i, 0 \leq s \leq n] = P[J_{n+1} = j, T_{n+1} - T_n = t | J_n = i] : = q_{ij}(t). \]

then we recover the classical semi-Markov chain model.

3 Application to real data

To check the validity of our model we perform a comparison of the behavior of real data and wind speeds generated through Monte Carlo simulations based on the models described above. In this section we describe the database of real data used for the analysis, the method used to simulate synthetic wind speed time series and, at the end, we compare results from real and simulated data.

The data used in this analysis are freely available from http://www.lsi-lastem.it/meteo/page/dwnldata.aspx. The station of L.S.I. -Lastem is situated in Italy at N 45 28' 14,9" - E 9 22' 19,9" and at 107 m of altitude. The station use a combined speed-direction anemometer at 22 m above the ground. It has a measurement range that goes from 0 to 60 m/s, a threshold of 0,38 m/s and a resolution of 0,05 m/s. The station processes the speed every 10 minute in a time interval ranging from 25/10/2006 to 28/06/2011. During the 10 minutes are performed 31 sampling which are then averaged in the time interval. In this work, we use the sampled data that represents the average of the modulus of the wind speed (m/s) without considered a specific direction. The database is then composed of about 230 thousands wind speed measures ranging from 0 to 16 m/s. The time series, together with its empirical probability density distribution are represented in Figure 1.

To be able to model the wind speed as a semi-Markov process the state space of wind speed has been discretized. In the example shown in this work we discretized wind speed into 7 states chosen to cover all the wind speed distribution.

From the discretized wind speeds we estimated the probabilities \( P \) and \( G \) to generate synthetic trajectories for three semi-Markov models: a simple semi-Markov model of the first order named semi-Markov I, semi-Markov II a second order semi-Markov model in state (as described in section 2) and the second order semi-Markov model in state and duration is named semi-Markov III.

For comparison reason, we also generated a synthetic trajectory which follows a simple Markov model with transition probability matrix estimated
from the real data. We then ended up with five trajectories: one representing real data, three representing the three trajectories according to the three semi-Markov models and the last one a Markov chain. The three time series are used in the following to compare results on the time lagged autocorrelation function. Real data are, in fact, strongly autocorrelated and the autocorrelation function decreases rapidly with time. We then tested our models to check whether they are able to reproduce such behavior.

If \( w \) indicates wind speed, the time lagged (\( \tau \)) autocorrelation of wind speed is defined as

\[
\Sigma(\tau) = \frac{\text{Cov}(w(t + \tau), w(t))}{\text{Var}(w(t))}
\] (18)

The time lag \( \tau \) was made to run from 1 minute up to 1000 minutes. Note that to be able to compare results for \( \Sigma(\tau) \) each simulated time series was generated with the same length as real data. Results shown in Figure 2 indicate that semi-Markov models reproduce better the autocorrelation present in real data especially if a second order semi-Markov chain is used to generate synthetic data. We can also notice that the second order semi-Markov model in state and duration (Model III) has to be preferred to the second order semi-Markov model in state (Model II) because it exhibits a slightly higher autocorrelation uniformly in time and also because, asymptotically, its autocorrelation has the same slope of that observed in real data whereas in the Model II, the autocorrelation drops to zero. Although the evidence shows the semi-Markovian nature of the studied phenomenon, probably a third/fourth order semi-Markov chain would be needed to decrease the difference between autocorrelation of real and simulated data. In our view this approach would be too much computationally and data consuming and further research on a simplified, but still with longer memory, model is needed.
4 Conclusions

Wind speed is a stochastic process for which a completely satisfactory model is still lacking. Many authors have used Markov chain to model the process but this approach does not give the same persistence present in real data. We presented in this work three semi-Markov models with the aim of generate synthetic wind speed data. We have showed that all our models perform better than a simple Markov chain in reproducing the statistical properties of wind speed data. In particular, the model that we recognized as being the more suitable is our third model which is a second order semi-Markov process in state and duration. We conclude that semi-Markov models should be used when dealing with wind speed data.

References

Insect efficient progeny distribution and demographic entropy

Petros Damos\textsuperscript{1,2}

\textsuperscript{1} Aristotle University of Thessaloniki, Department of Plant Protection, Laboratory of Applied Zoology and Parasitology, Greece
E-mail: damos@agro.auth.gr
\textsuperscript{2} Tech. Edu. Inst. Western Macedonia, Department of Agricultural Technology, Laboratory of Plant Protection, E-mail: petrosdamos@gmail.com

Abstract: Evolutionary ecology, in its broadest sense, involves the integration of two distinct disciplines: the description of genetic change within a population due to mutation and natural selection and the dynamics of age distribution and population sizes as a consequence of the interaction of an organism with its environment. This work is focused on the second discipline and makes efforts to use demographic entropy to evaluate the capacity of insect species to survive and reproduce under different environmental regimes. Demographic entropy is an analog of the Gibbs–Boltzmann entropy in statistical mechanics and is used as a highly qualitative predictor of species net reproductive success. Based on theoretical and empirical considerations the relation between demographic entropy and the reproductive success under different abiotic conditions is examined for representative insect species. Differences on the major demographic parameters as well as on their related Entropies can reflect the observed resource allocation behavior and microhabitat selection by the species. Moreover, in order to capture the uncertainty in the age of the mother of a randomly chosen newborn, and how it is affected after a finite number of time steps a population theoretical framework is presented in which life tables are treated as Leslie Matrices in a Markov process. Starting with non-equilibrium states an example of a finite chain sequence is generated to identify the distance to the equilibrium states for extreme condition. Hence in contrast to worm blood animals, cold-blood species and insects in particular have the ability to greatly alter their metabolic rates, as well as their related demographic parameters, in response to environmental factors. In this context, the simple reformulation of the stable age distribution as convergence to the stationary probability distribution of a Markov chain is able to reveal equilibrium states plasticity due to environmental factors and thus define the optimum conditions for successful progeny distribution based on classic ergotic theory.

Key words: Insect Ecology, Bio-Demography, Life Tables, Leslie Matrix, Information theory, Markovian Entropy, ergodicity

Introduction

In this work we are mostly interested in describing age transitions of poikilothermic organisms (e.g. insects and related arthropods) in terms of demographic entropy. The interest is to describe the degree of which population entropy is affected by environmental conditions and whether this
affects the age structured population transitions rate to the stable age distribution.

Since life table analysis and demography involves the modelling of time to event data, such as birth, death and survivorship and these times are subjected to random variations, one can use certain probability schemes to describe unpredictability of the system (e.g. population cohort) and age failure. As in the case of statistical quantum mechanics and machine failure theory, in which information entropy is related to the uncertainty of the realisation of a random variable, demographic entropy is related to the average probability of the species survivorship function.

Within the aims of the current work, building on preview reviews, is to provide a simple account for theoretical population ecologist and field experimentalist. Efforts are made to integrate fundamental theoretical results of demographic entropy with examples accrue from insect population studies are also presented.

**Bio-demography and species specific survivorship budged**

Bio-demography is grounded on the main principles and aspects of demography but is extending in the fields of ecology and includes the study of population sizes (number of individual within population), organism distribution arrangement (in time and space scales), population structure (distribution by age and sex) and population variations (temporal evolution of total growth and decline due to several factors. The general life table structures and processes that shape them are further used to draw scientific inferences upon the species specific population potential (Shryock et al. 1976). Based on life tables on can further proceed to the definition of the *survivorship budged* of particular species under certain conditions. The *survivorship budged* reflects a qualitative criterion that encompasses the capacity of increase of a given species under certain conditions.

The *survivorship budged* of a given species allows as to estimate, among other demographic statistics, the net reproductive rate per generation, the intrinsic rate of population increase and the Fisher’s reproductive value, which constitutes a relative measure of the contribution of the different ages to the ancestry of future generations (Gutierrez 1996). The *survivorship budged* normally follows the fate of a large cohort of individuals and has the following components or demographic parameters:

- The net reproductive rate ($R_0$), which equals according to the approximate (discrete method):

\[
R_0 = \sum_{x=a}^{b} l_x \cdot m_x
\]

or, by the exact method:

\[
R_0 = \int_{0}^{a} l_x \cdot m_x \, dx
\]
\( \beta \). The intrinsic rate of increase \((r_0)\):

\[
r_0 = \frac{\ln(R_0)}{T},
\]

To date, the intrinsic rate of increase was first introduced by Dublin and Lotka (1925) (Dyblin and Lotka 1925) and because it is the most inclusive demographic parameter, by the sense that it actual represents the species fitness (referred as demographic fitness), it is often used in most demographic studies.

\( \gamma \). The birth rate \((b)\):

\[
b = \frac{1}{\sum_{x=1}^{a} L_x \cdot e^{-r \cdot x}}
\]

\( \delta \). The death rate \((d)\):

\[
d = b - r
\]

\( \epsilon \). The finite capacity of increase \((\lambda)\):

\[
\lambda = e^r
\]

\( \zeta \). The mean generation time \((G)\):

\[
T = \frac{\sum_{x=d}^{b} x \cdot L_x \cdot m_x}{\sum_{x=d}^{b} L_x \cdot m_x}
\]

\( \eta \). The doubling time \((DT)\):

\[
DT = \frac{\ln(2)}{T}
\]

\( \iota \). The life expectancy \((e_x)\):

\[
e_0 = \sum_{x=0}^{a} L_x
\]

or

\[
e_0 = \int_{0}^{\infty} l_x \, dx
\]

\( \theta \). Life table entropy (by using life table notation instead of the information theoretic approach that is later described) is:

\[
H = \frac{\sum_{x=0}^{a} e_x \cdot d_x}{e_0}
\]

where, \(l_x\): number of individuals life at time \(x\) (age specific survival), \(m_x\): number of female offspring/female at time \(x\) (age specific fertility), \(L_x\):
number of days alive at time \( x \) and \( \omega \) age of last reproduction (Lee 1992, Carey 1993, Carey 2001). The parameter can be also calculated by the derivation of the discrete version of the Lotka equation:

\[
l = \sum_{x=0}^{\infty} e^{-r(x+1)} l_x m_x
\]


**The Entropy formalism and demographic heterogeneity**

In principle population entropy is based upon the conceptual framework of Shannon’s statistical information (Shannon 1948, Shannon and Weaver 1949, Reza 1994). Information entropy, more precisely the Boltzmann-Gibbs-Shannon entropy, for discrete and continuous variables, is by definition the quantity (Papoulis 1991, Papoulis and Pillai 2002, Antoniou et al. 2004, Damos et al. 2011b).

\[
H(X_a) = -\sum_{i=1}^{n} p_i \ln p_i
\]

\[
E \{- \ln f(x)\} = \int_{-\infty}^{\infty} f(x) \ln f(x) dx,
\]

where \( p_i \) is the probability mass function of outcome \( x_i \) and \( f \) the probability density function, while \( p(x) \) is a function defined only for \( x=x_i \) and such that \( p(x_i)=p_i \).

The above measure can be interpreted as the uncertainty about the variable \( X_a \) and equals the information gained when it is observed. Above equation (13) is defined according to the natural logarithm of the probability, instead of the logarithm having base 2, as originally defined by Shannon in channel communication (Reza 1994):

\[
H(X_a) = -\sum_{i=1}^{n} p_i \log_2 p_i
\]

Moreover, entropy can be extended also to the case of continuous random variables having probability density functions of the general form (Damos et al. 2011b):

\[
Pr[a < X \leq b] = \int_{a}^{b} f(x) dx, \forall x \in R,
\]

and after a definition of a finite number classes. The respective cumulative density function \( F \) is:

\[
F(x) = \int_{-\infty}^{x} f(t) dt = \int_{a}^{x} f(t) dt
\]

when \( f \) is defined for \( t > a \).
Demetrius in (1974) introduced the concept of population entropy $H$ as a summary statistic arguing that it is significant as $r$ to define species fitness. In practical terms the maximization of population entropy $H$ under various constraints yields to distributions of reproduction and survivorship (Demetrius, 1974, 1977, 1978, 1983, Demetrius and Ziehe 1984). Thus by considering that $l(x)$ represents the probability that an individual born at age zero survives to age $x$, and that $m(x)$ is the age-specific fecundity of cohort life table on equations the population entropy can be derived (Demetrius 1978) as:

$$H = -\int_{0}^{\infty} q(x) \log q(x) \, dx$$

where,

$$q(x) = \frac{l_x m_x}{R_0}$$

is the probability density function of the age of reproducing individuals. If $V_x=l_x m_x$ is the net maternity function and for fixed $m_x=1$, then

$$H = -\int_{0}^{\infty} \frac{l_x}{e_0} \log \frac{l_x}{e_0} \, dx = -\int_{0}^{\infty} \log \frac{l_x}{e_0} \, dx + \log e_0$$

and

$$H^* = -\frac{\int_{0}^{\infty} \log l_x \, dx}{\int_{0}^{\infty} l_x \, dx}$$

where $H^*$ normalised entropy. The above equation is an analogue of the Boltzmann–Gibbs definition of the entropy of a thermodynamic system. The numerator is analogous to the Shannon–Weaver measure of the amount of information in a message; while in the demography context it measures the variability of the contribution of different age classes to the stationary population.

Nevertheless, in this work classical demographic entropy is used as criterion to measure heterogeneity of insect population in relation to different temperature regimes, while later Markovian entropy is used to describe population robustness and convergence rate to the stable age distribution. Figure 1 for instance depicts the effect of five temperature regimes upon the population entropy of a representative moth species (i.e. *Anarsia lineatella* Lepidoptera: Gelechiidae).
Figure 1 Effect of five temperatures (°C) on the demographic entropy of a poikilothermic organism (i.e. insects and related arthropods). Values impose constraints on the shape of the net-reproductive function.

It is obvious that demographic heterogeneity is lower in the intermediate rearing temperatures but considerably higher at the extremes. Values are ‘snapshots’ of demographic entropy at an initial time step and can be interpreted in terms of how fitness (e.g. reproduction strategy) is modified due to alterations in temperature regimes.

From an information theoretical standpoint, where entropy is referred to as a measure of information, minimum $H$ yields minimum uncertainty (high information and population stability), and while maximum $H$ represents maximum uncertainty (low information and population stability).

**Population convergence to the stable age distribution**

The population entropy introduced by Demetrius is shown to have a precise dynamical meaning as a measure of convergence rate to the stable age distribution (Tuljapulkar 1982a,b) and related projections of population growth on the basis of survival and fertility assumptions date back to 1895 by Cannan. Whelpton in 1936 introduced matrix methods in population studies, while in the early 40’s, Bernardelli (1941), Lewis (1942), and Leslie (1940,1945) successfully formalised the methods to project populations. Here a brief account of the Leslie model is made and relative analogies to the Markov chain models are in short described.

Let $n_x$ be the number of organism (e.g. insect) in age $x$ (or stage for stage structured models) at time $\tau$, $S_x$ is the survival of the cohort in age interval $x$ to $x+1$ and $m_x$ corresponds to the average number of female offspring per female in age interval $x$ to $x+1$. These amounts are related as follows (Gutierrez 1996, Kot 2001):
\[ n_{x+1,r+1} = n_{x,r}S_x \]

and

\[ n_{0,r+1} = \sum_{x=0}^{n} n_{x,r}m_x = n_{0,r}m_0 + n_{1,r}m_1 + n_{2,r}m_2 + \ldots \]

The above equations can be combined according to the following projection scheme (Kot 2001):

\[
\begin{pmatrix}
m_0 & m_0 & m_1 & \ldots & m_{i-1} \\
m_0 & 0 & 0 & \ldots & 0 \\
S_0 & 0 & 0 & \ldots & 0 \\
0 & S_1 & 0 & \ldots & 0 \\
\vdots & & & & \\
0 & 0 & 0 & \ldots & 0 \\
\end{pmatrix}
\begin{pmatrix}
n_0 \\
n_1 \\
n_2 \\
\vdots \\
n_{i-1} \\
\end{pmatrix}
= \begin{pmatrix}
n_0 \\
n_1 \\
n_2 \\
\vdots \\
n_{i-1} \\
\end{pmatrix}_{r+1}
\]

It is convenient to denote by upper case bold faced symbols as Matrices (e.g., \( M \)) and vectors (e.g. \( N \); vectors are column vectors by default). The superscript \( \tau \) denotes the transpose. The symbol \( \text{diag}(x) \) denotes the matrix with the vector \( S \) on the diagonal and zeros elsewhere. The vector \( e \) is a vector of ones, and the vector \( e_i \) is the \( i \)th unit vector (i.e., the vector with a 1 in the \( i \)th location and zeros elsewhere that consist an adjacency matrix).

Thus, in terms of matrices the relative analogue is

\[ MN_{\tau} = N_{\tau+1} \]

and \( L \) is the projection matrix or Leslie matrix. The vector of age-class abundances in any given time point, to relative analogies as in the case of the limiting probability of Markov process, is:

\[ MN_{\tau} = N_{\tau+1} \]

The similarities of the above equation to that of a Markov process as it will be shown in the next section are obvious.

In order to find the stable age distribution the interest is to define a vector in which each class grows by the same factor \( \lambda \) at each time step:

\[ \lambda N_i = N_{(r+1)} \]

Here the second parts are equal and so the age distribution must thus satisfy:

\[ N_i \lambda = MN_i \]

or

\[ (M-I)N_i = 0, \]

where \( I \) the \( n \times n \) identity matrix and is \( M \) considered as eigenvector, corresponding to eigenvalue \( \lambda \). For the eigenvalue \( \lambda \) it holds \( \text{rank}(L-I)<k \), or \( \det(M-I)=0 \)

Solving for the eigenvalues and eigenvectors of the Leslie matrix for instance by direct solve to the homogenous set of equations of \((M-I)N_i=0\) for non-trivial solution considering the left-hand side matrix be singular, and that the determinant of the matrix satisfies:


\[ |M - \lambda I| = 0. \]

Moreover, the characteristic polynomial of \( M \) is:

\[ p_M(\lambda) = \det(M - \lambda I). \]

If we further denote \( Q \) the matrix having as columns the eigenvectors \( u_1, u_2, \ldots, u_n \) corresponding to eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \). It holds:

\[ Q^T = Q^{-1} \]

\[ Q^T M Q = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \]

and

\[ \det(M - \lambda I) = \det(Q^T) \det(M - \lambda I) \det(Q^T) = \det(Q^T A Q - \lambda I) = 0. \]

Considering that the Leslie matrix constitutes a directed graph one can now proceed to the estimation of the spectrum of \( M \) which is the set of the eigenvalues of \( M \) \( \lambda \) roots of \( p_M(\lambda) \).

To date, eigenvalues and eigenvectors are usually used to study the change in a population over time in a dynamical system and to give a meaningful biological interpretation. Thus, they can be used to demonstrate whether a population is increasing, decreasing or staying constant under certain conditions. Thus if \( \lambda = 1 \) the population is stationary, if \( \lambda > 1 \) population increases and overpopulation is experienced, while if \( \lambda < 1 \) the population decreases and progressively diminishes, while the logarithm of the dominant eigenvalue:

\[ r = \log(\lambda) \]

results to the annual rate of increase.

In principle the role of the eigenvectors is to provide the stable age distribution between age classes of the populations represented by the right and left eigenvectors. It should be noted that the left eigenvector of a matrix is distinct from its right. If \( u \) denotes the eigenvectors and * the complex-conjugate transpose by solving the equation:

\[ \nu^* M = \lambda u^* \]

we can determine the stable vector, or the eigenvectors \( u^* \).

Note that the eigenvector corresponding to the eigenvalue \( \lambda = 1 \) is referred to as the stable age distribution for a specific population. Figure 2 is representative example which depicts the projections of the Leslie matrix model for the population growth of an insect (i.e. lepidopteron) that was reared at temperatures close to its upper extreme threshold (30°C).
Let now $M=\mu_\ell$ be a $k \times k$ Leslie Matrix as defined above having a first row of reproduction, $\mu_1=m_0$, subdiagonal of survival rates $S_{i-1,j}=p_i$ and all other elements are zero. Assuming $M$ irreducible and primitive that has dominant eigenvalue $\lambda_0$, right eigenvector $u$ and left eigenvector $v$. For any initial vector $N_0$ of population numbers the population vector $t$ time intervals later is:

$$N_t=M^tN_0 \quad \text{and} \quad \lambda^tN_t \rightarrow (v^* N_0)u^*$$

Thus on the basis on the same survival and fertility life tables data, populations convergence to a stable age distribution in which the species cannot sustain survivorship and reproduction and the population progressively decrease towards extinction (Figure 3).
Figure 3 A capture history representing the fate of an insect population over the study period under an extreme temperature (31±1°C, 16:8L:D, 65±5%RH). It is clear that the population is progressively decreases towards extinction.

Since the Leslie model can be regarded as a modification of the Markov chain model that treats the demographic categories of the projection system as states of the state space (i.e. elements of the state vector) one can further proceed to the estimation of entropy of the above stochastic process. In particular considering the process of age transitions as stochastic Markov chains on can estimate the conditional probabilities easily based on the following property:

\[ P[X_{n+1} = x_{n+1} | X_n = x_n, ..., X_1 = x_1] = P[X_{n+1} = x_{n+1} | X_n = x_n] \]

where, \( X_n \) is a random process on a probability space \((\Omega, F, P)\), the random process, will be finite (or infinite) sequences \( \{x_i\} \) of random variables taking values in a discrete (countable or finite) set \( S \), the elements of which are called states. One can further define the one step transition probability from state \( i \) to state \( j \) at instant \( n \) as follows:

\[ p_{ij} = P[X_n = j | X_{n-1} = i] \]

and if the Markov chain satisfies the homogeneity property, this implies that one step transition probability \( p_{ij} \) while be independent of \( n \) (i.e. the instant when the transition actually occurs). Nevertheless, since the Leslie model is not row stochastic a ‘normalised transformation’ should be first made. This can be made as described by (Tuljapurkar 1982):
If $U$ and $V$ are diagonal matrices in terms of the eigenvalue components, a stochastic description of Age structure can be made as follows (Tuljapurkar 1982):

$$U = \text{Diag}[u(1), u(2), \ldots, u(k)]$$
$$V = \text{Diag}[v(1), v(2), \ldots, v(k)].$$

By transposing both sides in $MN^\tau = N^{\tau+1}$, multiply by $V$ and divide by $\lambda^\tau$ we obtain:

$$(N^\tau V/\lambda^\tau)(V^{-1}M^*V/\lambda) = (N^{\tau+1}V/\lambda)$$

and for initial $N_0$ the normalized vector:

$$\pi_i^0 = \frac{N_0 V_i}{\sum N_0 V_j},$$

with:

$$\sum \pi_i^0 = 1.$$

The Leslie model can be replaced by the equivalent forward Markov Chain having transition matrix $P_F$:

$$\pi_i^{(n)} = \pi_i^{(n-1)} P_F,$$

where:

$$\pi_i^{(\tau)} = \frac{V N^\tau}{\lambda^\tau (V N_0)}$$

and

$$P_F = (1/\lambda) V^{-1} M V$$

The above matrix is row stochastic:

$$\sum a_{ij} = 1, \forall i \in \mathcal{S},$$

while the demographic interpretation of the vector $\pi_i^\tau$ is determined by recalling that the elements of the eigenvectors correspond to reproductive values. The component $\pi_i^\tau$ equals the reproductive value of individual’s on age $i$ at time $\tau$, divided by the total reproductive value of the population at time 0, and discounted by $\lambda^\tau$.

In terms of the Leslie matrix and a Markov related process can be also defined in respect to the eigenvalues $U$ by:

$$P_B = (1/\lambda) U^{-1} M U$$

In this formula eigenvalues $P_F$ and $P_B$ are identical; the initial states of the process with transition matrix $P_B$ cannot be naturally identified (i.e. backward process). For more details refer to (Feller 1968).

One can now proceed to estimation of the convergence rates and entropy estimation towards equilibrium as in the case of physical problems addressed by erotic theory.

For a given Markov chain with transition probability matrix $M$ and state space $S = \{0, 1, 2, \ldots, k\}$ the initial probability of state $i$ is given by:

$$\pi_i^{(0)} = P(X_0 = i), i \in S.$$ It can be shown (Cox and Miller 1965) that:

$$\pi_i^{(n)} = \pi_i^{(n-1)} M$$
Hence $\pi_i^{(n)} = \pi_i^{(n-1)} M = \ldots = \pi_i^{(0)} M^{(0)}$, and if there is a steady state as $n \to \infty$, then the stationary distribution is governed by the evolution of successive powers of the transition matrix. If the parameter space $T$ is continuous then the transition probabilities from state $i$ to state $j$ at time $\tau$ are defined:

$$p_{ij}(\tau) = P(X_{\tau} = j | X_{\tau-1} = i), \quad \forall i, j \in S \text{ and } t \geq 0.$$ 

In Figure 3 a Markov chain probability transition model of three states (i.e. stages S1, S2, S3).

![Figure 3](image-url)

**Figure 3** Generated Markov chain model that corresponds to an initial transition probability matrix scheme having three states (i.e. three population stages) and convergences to the limiting probability. $P^n = p_{ij}(t)$ is irreducible and aperiodic so that $p_{ij}(\tau) \to \pi$ as $\tau \to \infty$ where $\pi_i^{(n)} = \pi_i^{(n-1)} M$.

With the above probability scheme, one can further estimate the Markovian entropy $H$, which equals the average uncertainty of the system for moving one step ahead when starting with one given initial state $S_i$:

$$H_i = -\sum_{j=1}^{n} p_{ij} \log p_{ij}$$

More generally, the set of events of going from $S_i$ to any other state in $r$ steps constitutes a finite complete probability scheme with entropy:

$$H_i = -\sum_{j=1}^{n} p_{ij}^r \log p_{ij}^r$$

where $p_{ij}$ stands for the probability of any one of a discrete chain moving from the $i$th to the $j$th state in $r$ steps. Thus, the entropy of the chain for
moving $r$ steps ahead from the initial states when the initial probabilities are specified is:

$$H' = \bar{H}' = -\sum_{i=1}^{n} p_i H_i' = -\sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_{ij}^r \log p_{ij}^r$$

For more details refer to Reza (1994), Papoulis and Pillai (2002). The average entropy that corresponds to the Markov chain model of Figure 4 is generated in Figure 5. As it can be seen the stochastic process is ergotic and limiting equilibrium distribution in which entropy convergence after a number of time steps.

![Entropy convergence graph](image_url)

**Figure 5** Entropy convergences of a typical stochastic process as illustrated in Figure 3 for three population stages. Entropy reaches equilibrium after a finite number of time steps.

Finally it should be made clear that if $\pi_i (\omega) \rightarrow \omega$, population entropy, or Kolmogorov-Sinai entropy (Kolmogorov 1959, Cornfeld et al. 1981), in terms of the transition matrix elements $\pi_i$ and the stationary distribution can be also represented as:

$$H = -\sum_{i,j} \omega(i) p_{ij} \ln p_{ij}, i,j=1,2,\ldots,k,$$

By following the transformation proposed by (Tuljapurkar 1982) and by considering that the mean generation time $T$ is:

$$T = \sum_{i} (\lambda^{-1} \phi_i)$$

the Leslie matrix reduces to:
\[ H = -\frac{1}{T} \sum_{i} (\lambda^{-1} \phi_i) \ln(\lambda^{-1} \phi_i), \]

\( H \) determines the asymptotic rate at which this convergence occurs. Detailed derivation of the above mathematical function, as well as the basic assumptions and modifications of the original formula, are covered in details in Tuljapurkar (1982a,b, 1990), Tuljapurkar et al. (2003) and references cited.

Finally, the calculation of entropy convergence towards the stable age distribution can be easily made by introducing a simple norm, which estimates the distance \( (D_i) \) from any initial state \( (H_{\tau i}) \) to the equilibrium \( (H_{eq}) \):

\[ \left\| H_{\tau i} - H_{eq} \right\| = D_i \]

The above norm represents a “velocity” measure of convergence towards equilibrium. In addition, the Kullback–Leibler distance \( (KLD) \) (referred also as information divergence or relative entropy), has been also proposed in demography to study convergence to a stable age distribution (Tuljapurkar 1982b, Al-Khafaji et al. 2007). The \( KLD \) is by definition a directional measure of the distance between two probability distributions (Kullback and Leiber 1951, Burnham & Anderson 1998) and for discrete probability distributions (Al-Khafji 2007) is:

\[ KLD(p, \pi) = \sum_{i=1}^{k} p_i \log\left(\frac{p_i}{\pi_i}\right) \]

where \( k \) is the number of possible outcomes, \( p \) is the reference distribution and \( \pi \) is the approximating distribution. The advantage of the \( KLD \) is that it provides substantially greater information about age-structure than a distance based on a norm or on Euclidean distance. For more details of concerning the demographic analogues of the given distance measures refer to Tuljapurkar (1982a,b, 1990) and Al-Khafji (2007).

**Summary and Conclusions**

In this work entropy is used as a measure of population heterogeneity for insects and related arthropods and how in relation to different temperature regimes. Such kind of information can further used as qualitative indicator to infer on population robustness (i.e. population disorder and intractability).

First some classical examples of temperature effects on the insect survivorship budget in terms of life table estimated demographic entropy are given, while later we have mostly focused on the theoretical description of population age transitions to the stable age distribution and representative examples are presented having relative analogies to Markov processes.

Recent activity in demographic modelling is aimed mainly at a better definition of the population system to be modelled in to representing more accurately the population dynamics and dependence to certain conditions. Algebraic demographic modelling in particular combines the modellers’
interest in the age dimension of a population. This is important considering that demographic responses (e.g. birth and death rates) in a population are not all identical but age dependent (Foerster 1959, Rosa et al. 2011, Damos and Savopoulou-Soultani 2012). Therefore life table analysis, taking into account the age dependent mortality and reproduction in the estimation of the demographic parameters, is more suitable for the description population growth closer to reality compared to simple reproductive summary statistics (Carey 1993).

To address how demographic and evolutionary constraints limit the persistence, Demetrius in (1974) introduced the concept of population entropy $H$ as a summary statistic arguing that it is significant as $r$ to define species fitness. In fact the intrinsic rate of increase and population entropy summarizes different biologically significant consequences of age-specific schedules of mortality and fecundity. One advantage of population entropy over the intrinsic rate of increase is that it is related with the rate of decay of population fluctuations and thereby provides a precise measure of demographic stability (Ziehe and Demetrius 2005). This property encloses information upon the ability of population to sustain perturbations induced by intrinsic or extrinsic factors and can be used as measure of population robustness. Based on the Demetrius entropy conceptual framework, the species are divided in two broad categories; considered either as opportunistic, characterized by a unidirected increase of entropy, or as opportunistic, displaying an unidirectional decrease in entropy for large populations and random changes in entropy for small populations (Kowalt and Demetrius 2009).

From a practical standpoint, demographic properties of a population do constrain to the rate of which species adapt to human disturbed environments. Repeated use of pesticides for instance can cause undesirable changes in the gene pool leading of a species due to artificial selection. Through this process, the population gradually develops resistance to the pesticide and detection of resistance individuals usually show fitness disadvantages in the absence of the artificial selection factor (McKenzie et al. 1982, Roush and Daly 1990, Herron and Rophail 1993).

Population entropy numerical studies should be thus profitable in insect ecology in cases in which we won’t to determine the rate of convergence to equilibrium in situations where several activities and factors disrupt initial population stability. This information is also important to predict colonization dynamics and/or species invading potentials to new areas and simulating life history evolution (Granett et al. 1983; Trichilo and Leigh 1985; Omer et al. 1992; Vargas et al. 1997).

Concluding, for a given insect species, fitness and reproductive performance is a function of the environment and the deviation of the mean demographic properties can measure how fast genetic shift occurs. Emanations of these transactions could be expressed through the evolution of new phenotypes and/or strains having variants life traits (Hoffmann and Blows 1994, Chown and Terblanche 2006, Chown et al. 2009, Whitman.
This plasticity is a direct impact of the environment on the development of successful phenotypes, which apart of morphological diversity involves also the development of physiological, behavioral and demographic responses (Chown and Terblanche 2006, Chown et al. 2009, Whitman 2009). The degree of phenotypic expression, here entropy related fitness measure, depends upon the species intrinsic factors, the time scale of exposition and the intensity of the selection factor. Under an evolutionary perspective the genetic variance as product of phenotypic variance and heritability, drives evolution and allows persistence of individuals to new conditions.

References
The effect of mortality models on life expectancy and other mortality indicators

A. Debón¹, S. Haberman², F. Montes³, and E. Otranto⁴

¹ Centro de Gestión de la Calidad y del Cambio, Universitat Politècnica de València, 46022 Valencia, Spain
(E-mail: andeau@eio.upv.es)
² Cass Business School, City University, EC2 London, United Kingdom
(E-mail: S.Haberman@city.ac.uk)
³ Departament de Estadística i I.O., Universitat de València, 46100 Burjassot, Valencia, Spain
(E-mail: montes@uv.es)
⁴ Departament de Estadística i I.O., Universit degli Studi di Messina, Piazza Pugiatti, 89122 Messina, Italy
(E-mail: eotranto@unime.it)

Abstract. The parametric model introduced by Lee and Carter in 1992 for projecting mortality rates in the US has been a seminal development and has been widely used since then. Different versions of the model, incorporating constraints on the data, and different adjustment methods have led to improvement. All of these changes have increased the complexity of the model with a corresponding improvement in goodness of fit, however, there is little change in the ability of forecast of life expectancy in comparison with the original Lee-Carter model, according to some authors.

To evaluate to what point the increments in the complexity and computational cost of the models are reflected in the forecast of such indices as life expectancy and modal age at death, among others, we have compared two different models: the original Lee-Carter with one and two temporal parameters. The three sets of predictions so obtained are compared using a mixture of block-bootstrap techniques and functional data analysis.

Keywords: Mortality indicators, Block-bootstrap, Functional Data Analysis.

1 Introduction

In the context of recent demographic changes, the development of new models for building life tables and their projection is presented as a point key research. Life expectancy reflects these changes but its effects are diminished due to its robustness. If, moreover, we bear in mind that life expectancy offers no information as to whether this improvement is the same for different age groups, it is important and necessary to turn to other mortality indicators whose past and future evolution in Spain we are going to study.

An appropriate set of indicators for the study of all these phenomena should include an indicator of infant mortality, the Lorenz curve, the Gini
index and the modal age at death. These indicators were applied to Spanish mortality data for the period 1981-2008 for the age range 0 to 99 in [3]. The main conclusions obtained were:

- mortality in Spain improved in both the observed period, 1981-2008, and the forecast period, 2009-2028,
- future improvement is expected to be more sustained than that experienced during the period observed,
- the evolution of the modal age at death, the Lorenz curve and Gini index also confirmed that Spanish mortality displays both expansion and rectangularization,
- the mortality of women is better than men, meaning longer life expectancy and modal age at death, and a lower Gini index.

All these indicators can be projected using the projections of $q_{xt}$, obtained from different methodologies. The errors associated with these estimations can be calculated by means of a block-bootstrap methodology [7] and a prediction interval can be provided. The aim of this paper is to evaluate whether different mortality models produce significant differences in the projections of some of these mortality indicators. The reason is that improvement in goodness-of-fit is achieved by increasing the complexity of the models proposed in the literature at the cost of a higher computer complexity and costs. We want to study if this improvement in goodness-of-fit will be reflected in significant differences between the indicators.

The study has a second aspect that we wish to point out: the method used to assess these differences. On the one hand is the use of block-sampling techniques to obtain bootstrap confidence intervals for mortality indicators. Some authors have drawn attention to the narrow intervals obtained by classical bootstrap techniques; however block-bootstrap techniques produce confidence intervals which are more realistic. The second methodological innovation is the use of functional analysis techniques to detect possible differences between the projections of the indicators obtained with the different models. The reasons for the use of functional data are two-fold: firstly, because the projections of one indicator over time are correlated values and proper analysis requires longitudinal data techniques or functional data analysis, as we propose; secondly, because so far the comparison has been carried out by comparing graphs visually, depriving the decision of objectivity.

The paper is structured as follows. Section 2 is devoted to describing Dynamic Life Tables and a brief summary of Lee-Carter models. Section 3 introduces the block-bootstrap techniques for building confidence intervals. Statistical techniques allowing the comparison of mortality indicators are presented in Section 4. Section 5 is devoted to the results of the analysis of Spanish mortality data by means of the above indicators. Finally, Section 7 establishes the conclusions to be drawn from the results in the previous section.
2 Lee-Carter models

We consider a set of crude mortality rates  $\dot{q}_{xt}$, for age $x \in [x_1, x_k]$ and calendar year $t \in [t_1, t_n]$, which we use to produce smoother estimates, $\hat{q}_{xt}$, of the true but unknown mortality probabilities $q_{xt}$. A crude rate at age $x$ and time $t$ is typically based on the corresponding number of deaths recorded, $d_{xt}$, relative to those initially exposed to risk, $E_{xt}$.

A dynamic life table is a rectangular mortality data array $(q_{xt})$, where $x$ denotes age and $t$ denotes calendar time. Each column in this array represents the constituents of the period life table for year $t$.

The Lee-Carter Model to the central mortality rates is well-described in [6]. We are going to consider the extended version of this model to logit death probability $q_{xt}$,

$$\ln \left( \frac{q_{xt}}{1 - q_{xt}} \right) = a_x + \sum_{i=1}^{r} b_i^x k_i^t + \epsilon_{xt}. \quad (1)$$

In our application to the Spanish data of mortality we have used (1) with $r = 1$ and $r = 2$, and consequently the corresponding models will be called $LC_1$ and $LC_2$, respectively.

The models are sufficiently well known and will not be considered further in this presentation. A detailed description of the model and its adjustment by different methods can be found in [4].

Forecasts for $q_{xt}$ with the Lee-Carter model are generated by first modelling $\hat{k}_t$ as a time series by using the Box-Jenkins methodology.

The set of indicators for this study are well-described in [3], life expectancy, the Gini index and the modal age at death. All these indicators can be projected using the projections of $q_{xt}$, obtained from an adequate methodology, which in our study is the family of Lee-Carter models. The errors associated with these estimations can be calculated by means of a block-bootstrap methodology [7] and, prediction intervals can be provided.

3 Block-bootstrap confidence intervals

Some authors argue that mortality predictions must be accompanied by measures of sensitivity and uncertainty. One way to combine all these sources of uncertainty is to use bootstrapping procedures. Therefore, we aim at obtaining prediction intervals for the mortality indicator by using a residual-based block-bootstrap, as [7] propose for deviance residuals, because this technique partially retains the underlying dependence structure in the residuals and generates more realistic resamples [5]. In this paper, block-bootstrap confidence intervals for the mortality indicator are going to be obtained using logit residuals and the binomial distribution assumption.
The procedure used is the following. Starting from the logit residuals, \( \hat{\varepsilon}_{xt} \) which are obtained from the original data,

\[
\hat{\varepsilon}_{xt} = \text{logit}(\hat{q}_{xt}) - \text{logit}(q_{xt}),
\]

ordered in a rectangular array \((\hat{\varepsilon}_{xt})\), where \( x \) denotes age and \( t \) denotes calendar time. To set up a new artificial set of residuals \( \hat{\varepsilon}_{nxt} \), we start with an empty rectangular array which has the same dimensions as the original matrix of residuals. The empty array is then partitioned into smaller rectangular blocks. Each block is replaced by a block of the same size, which is randomly selected from the original matrix. This block consists of all residuals in the rectangle to the southeast of the randomly selected element from the original matrix.

Estimated logit rates, \( \text{logit}(q_{xt}) \), are then set and the observed logit rates, for the \( n \)'th element of the sample, are obtained from the inverse expression

\[
\text{logit}(q_{xt})^n = \text{logit}(\hat{q}_{xt}) + \hat{\varepsilon}_{nxt}.
\]

With these new sampled logit rates, a new adjustment of the model is obtained which provides new estimations of the parameters. The process is repeated for the \( N \) bootstrap samples, which in turn provide a sample of size \( N \) for the set of model parameters, and the \( k_i \)'s are then projected on the basis of an ARIMA model, obtaining predictions for the mortality rates and the corresponding life expectancy and mortality indicators for the desired future years. The prediction intervals are obtained from the percentiles, \( IC_{95} = [p_{0.025}, p_{0.975}] \).

4 An ANOVA test for functional data

The experimental design we have carried out in order to test whether different modeling mortality ratios produce significant differences in the projections of mortality indicators is a two-way design, whose structure is shown in Table 1. It is a balanced design with the same number of repetitions in each cell, \( n_B \), exactly the number of block-bootstrap samples. Each one of these repetitions is a set of 20 values, the projected mortality indicator corresponding to the years 2011 to 2030. The factor model has two categories, the Lee-Carter with one time parameter, \( LC1 \), and the Lee-Carter with two time parameters, \( LC2 \). In turn, the factor residual also has two categories reflecting the origin of the residuals used in the bootstrap process, \( RLC1 \) and \( RLC2 \), according to whether they were obtained from the adjustment of original data with the \( LC1 \) or \( LC2 \) model.

As mentioned at the beginning, our aim is to check differences among the projections obtained with different models and different residuals. Therefore, we will resort to the method proposed by [1] which is based on the analysis of randomly chosen one-dimensional projections of functional data.
This method can perform a functional two way ANOVA with interactions. Following [1] we can write,

\[ X_{i}^{mod,res}(t) = m(t) + f^{mod}(t) + g^{res}(t) + h^{mod,res}(t) + \varepsilon_{i}^{mod,res}(t), \]  

where \( m \) is non-random and describes the overall shape of the projections, \( i = 1, \ldots, n_{B} \), and the functions \( f^{mod}, g^{res} \) and \( h^{mod,res} \) account for the main effect and interaction of model and residual. Finally, \( \varepsilon_{i}^{mod,res} \) are independent random trajectories centered on the mean.

### 5 Analysis of Mortality Data from Spain

The data used in this analysis come from the Spanish National Institute of Statistics (INE) (see their official web site at http://www.ine.es). In particular, we have worked with published life tables.

The models described in Section B with the expression \( r = 1 \) and \( r = 2 \) have been used to adjust mortality data in Spain for the period 1991-2010 and a range of ages from 0 to 99. The estimation of the parameters is carried out by means of maximum likelihood methods using the gnm library [9] written in the R language [8], as [2] propose. The adjustments have been made only for men and for the sake of brevity are not reproduced here. The high number of parameters estimated in the LC1 model, \( 100 \times 2 + 20 = 220 \) and the LC2 model, \( 100 \times 3 + 20 \times 2 = 340 \), cannot be fully presented in a paper of this length. A detailed description of the results of the models applied to Spanish mortality data can be found in [4].

Figure 1 is used to show the underlying dependence structure in the logit residuals for both models, so we are going to use block-bootstrap [7]. Now our main problem is how to select the block size. In the absence of firm theoretical guidance, [7] plot a correlogram and a contour map of the original raw residuals and compare these with the resampled residuals. If these plots match reasonably well, this gives confidence in the choice of blocksize as it supports the fact that they have a similar underlying dependence structure.

With respect to the block sizes, our initial guesses are based on the dependence structure observed for raw residuals in Figure 1. Therefore, we have 50 block \((5 \times 4)\)-bootstrap samples obtained from logit LC1 residuals and the same from logit LC2 residuals, which we are going to call RLC1 and RLC2, reflecting the origin of the residuals used in the bootstrap process.
Fig. 1. Residuals for age-period for $LC_1$ (left) and $LC_2$ (right) for men.

and according to whether they were obtained from the adjustment of original data with the $LC_1$ or $LC_2$ model as we mentioned in Section 4.

6 Prediction intervals for mortality indicators on block-bootstrap

Life expectancy remains the most familiar measure of longevity among demographers, and although it reflects the changes in mortality over time, it does so in a smooth way due to its robustness. This is the reason why other indicators are studied in this paper: modal age at death and the Gini index. The Gini index is a compression or dispersion of the mortality measure and the other two forecasted indicators are measures of the central tendency of mortality, life expectancy and modal age of death. All these indicators share the advantage of summarizing information about mortality independently of the age structure.

Forecasted mortality indicators for the period 2011-2030 were carried out using the block-bootstrap technique described in Section 3. Figures 2, 3 and 4 summarize their behavior for both models ($LC_1$ and $LC_2$) and different samples ($RLC_1$ and $RLC_2$). As seen in Figures 2, 3 and 4 all the predicted measurements and confidence intervals for all the models and samples behave in a similar way, which makes choosing one model over others difficult. Figures 2, 3 and 4 show the estimations and the corresponding prediction intervals for the corresponding mortality indicators. Life expectancy and the modal age at death continue to increase, and the Gini index continues to decrease.

With regard to the width and symmetry of the prediction intervals for life expectancy (Figure 2), the first feature to highlight is narrowness and the symmetry for the combination $LC_1$ model and $RLC_1$ (Figure 2 left) and the extended width but greater symmetry for $LC_2$ and $RLC_2$ (Figure 2 right); the explanations being the interaction between them. We expect this comment to be valid for the other two indicators also. But the intervals obtained for modal age at death show wider and more irregular intervals (Figure 4), which are narrower and more symmetric for the combination $LC_1$ model and $RLC_2$ (Figure 4 left) or $LC_2$ and $RLC_2$ (Figure 4 right). In the case of the Gini index, note (Figure 3) the interval narrowness and great
symmetry for the combination LC1 model and RLC2 (Figure 3 left) and for LC2 and RLC2.

Fig. 2. Predicted values for life expectancy at birth for LC1 (left) and LC2 (right) for men.

Fig. 3. Predicted values for the Gini index for LC1 (left) and LC2 (right) for men.

Fig. 4. Predicted values for modal age at death for LC1 (left) and LC2 (right) for men.

The p-values for functional ANOVA are shown in Table 2. According to model 3 the following null hypotheses have been tested,

\[ H_{0}^{mod} : f^{LC1} = f^{LC2} = 0 \]
\[ H_{0}^{res} : g^{RLC1} = g^{RLC2} = 0 \]
\[ H_{0}^{mod,res} : h^{LC1,RLC1} = h^{LC2,RLC1} = h^{LC1,RLC2} = h^{LC1,RLC2} = 0, \]

that is, there is no influence of the main factors and no interaction between them. From these results the three null hypotheses are accepted.
<table>
<thead>
<tr>
<th>Indicator</th>
<th>Model Residual</th>
<th>Model × Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expectancy</td>
<td>0.8400</td>
<td>0.7500</td>
</tr>
<tr>
<td>Gini index</td>
<td>0.6800</td>
<td>0.7000</td>
</tr>
<tr>
<td>Modal age</td>
<td>0.7200</td>
<td>0.9200</td>
</tr>
</tbody>
</table>

Table 2. p-values for functional ANOVA obtained from 100 bootstrap samples

7 Conclusions

When compared with similar studies carried out using Spanish mortality data, our results on life expectancy are slightly higher than those obtained by [1], [2] and [3] and very similar to life expectancy published by INE (see their official web site at http://www.ine.es). The comparison is possible for other indicators with [3], because as we point out in the Introduction, we have not found similar studies for these indicators by other authors. In relation to the work of other authors, we should highlight one distinctive feature of the methodology presented here, which is the possibility of comparing the projections of different models with an objective criterion. In short, we propose statistical tools which provide a clear framework for supporting decisions in mortality modelling.

References

Survival betterment as competitive leverage in insurance sector: profitability analysis for a class of participating variable annuities

Emilia Di Lorenzo¹, Albina Orlando², and Marilena Sibillo³

¹ Department of Mathematics and Statistics, Via Cinthia Monte S. Angelo, University of Napoli "Federico II", 80126, Napoli, Italy (E-mail: diloremi@unina.it)
² CNR, Istituto per le Applicazioni del Calcolo Mauro Picone, via P. Castellino 111 80126 Napoli, Italy (E-mail: a.orlando@iac.cnr.it)
³ Department of Economics and Statistics, Campus Universitario, University of Salerno 84084, Fisciano (SA), Italy (E-mail: msibillo@unisa.it)

Abstract. Population aging keeps on creating growth opportunities for life insurance sector; in light of this opportunity, the topics discussed within the paper are focusing on the definition of contractual profiles involving life annuities in order to obtain a reasonable risk-profit sharing between the parties. In particular survival indexed participating annuities are studied, for which a profitability analysis is provided by means of suitable balance-sheet indexes.

JEL Codes: C53, G17, G22, G32

Keywords: Actuarial Return on Gain, variable annuities, longevity risk.

1 Demographic evolution outline and the actuarial products supply

The phenomenon of demographic change develops through several aspects, such as birth decrease and peoples aging, whose management has now become a topic of interest for the international community and constitutes a crucial challenge for todays society, even with respect to the topic of solidarity among generations. The topic of macro-economic impact of the evolutional phenomenon of survival has already been brought to the attention of the First World Assembly on Aging, which took place in Wien in 1982; subsequently, the debate kept going on in Madrid in 2002, during the Second World Assembly on Aging (see [9]); today, the issue is constantly being monitored by EU bodies, aiming at transforming a challenge into an opportunity (see [7]).

In particular, among the plethora of questions related to the debated topic, the consequences of population aging within the sectors of social security and public spending are being examined (see: [11]; [7]).

It is appropriate to remind that the European Commission, in its Third Demography Report, published in April 2011, managed to synthesize the
growing increase of aging on the basis of the following data: the percentage of the European population older than 65 has increased to 13.7% in 1990; to 17.4% in 2010; projections made for 2060 even estimate a percentage equal to 30%. Furthermore, the so-called "oldest old", namely those individuals who are older than 80, are meant to significantly increase by 2060. Systematic aging will affect the whole Europe, however some regions will particularly be interested, as explained by the studies operated by the Committee of the Regions, recording a variability range which will, in 2030, sway between 10.4% and 37.3% (see [1]). There is little doubt about the magnitude of the impact of this set of data on a socio-economic field.

For what concerns the topics discussed in the paper, it is undoubtedly interesting to focus particularly on the consequences in the insurance sector, from two different perspectives: first of all, in terms of food for thought concerning the actual possibilities of such an important sector and, secondly, in relation to the social relevance of instruments capable of improving life conditions of people of an advanced age.

For what concerns the current size of insurance sectors, instead, Swiss Re provided a precise mapping of insurance industry on a global scale, which allows for a clear understanding of the consequences of the aging phenomenon in relation to the demand for aimed products (see [10]). The aforementioned report, based upon the data collected in 78 countries, representing 98% of global premium collection, highlights the positive trend of insurance sector, regardless of the unavoidable criticalities related to investment returns; in fact, the volume of premiums has globally registered an increase by 2.7%, thus resulting into a sum, which, despite being inferior to those preceding the financial crisis, exceeds 4300 billion dollars. For what concerns life sector, instead, Swiss Re’s report shows how population aging keeps on and will keep on creating growth opportunities; this represents, then, a set of opportunities which are to be strongly defended if we think about the current financial condition and, mainly, about public debt crisis in Europe - extremely dangerous for actives - and the complex regulatory implementation of Solvency II. On the basis of what explained, then, it seems evident that among the many peculiarities characterizing the complex demographic change phenomenon, the evolontional dynamics of survival, characterized, during the last decades, as well known, by a certainly positive trend, do occupy an important position, due to the consequences deriving from it in an economic, social and cultural environment. The topics discussed within the paper are mainly focusing on a sector which is closely connected to the matter concerning the evolution of survival, namely the security-insurance one. In this sense, the aging phenomenon has been studied as an opportunity for development of a proper financial sector, with the positive consequence of a competitive leverage. Such line of inquiry is not, anyway, unrelated to social wellness improvement goals, since it concretely responds to the needs of a relevant segment of the population, currently under the spotlight at a European level.
In fact, the words stated by the European platform Age (see [1]), concerning senior consumers rights, are particularly meaningful: “to guarantee older people’s rights as consumers as prescribed in the EU legal framework and to help raise awareness regarding those rights and how to make better decisions as consumers”.

Therefore, on the basis of what pointed out earlier, the impact of innovative contractual architectures aimed at risk-profit sharing between the parties, framed within the picture represented by the scenario described earlier, has a multi-faceted relevance; in detail, they may provide useful instruments, which would bring advantage to economic growth, on one hand on the level of insurance industry development (which now finds itself forced to face a growing demand of contractual typologies, appropriate to evolutional dynamics of survival; in particular, one goal stands above the others: giving an appropriate response to the increasing pension sector demands), on the other hand, on the level of safeguard for those who ask for investment instruments and possibilities fitting an elderly age. The afore mentioned advantages have to be framed within the development of competitive knowledge in a financial and actuarial field, where such a development undoubtedly implies a conscious use of the instruments proposed by financial and insurance engineering, which is largely fuelled by correct knowledge and scientific innovation.

2 Contractual architecture within a risk-profit sharing perspective

In the following we consider profit participating life annuities with installments scaled by a demographic index structured on the forecasted survival trend and the observed one. According to the contractual scheme introduced by Denuit et al. ([5]), the installments for to each insured aged \( x \) at issue, if alive at the payment time \( t \), are scaled by the factor:

\[
SI_{x,t} = \frac{x_{p}^{tech}}{x_{p}^{proj}}, \tag{1}
\]

where \( x_{p}^{tech} \) represents the survival probability for an insured aged \( x \) to age \( x + t \) inferable from a proper technical bases and \( x_{p}^{proj} \) the survival probability inferable from a projected demographic model.

Following Denuit et al. ([5]), we also introduce a minimum and a maximum threshold for the survival index, so overcoming marketing problems arising from increasing projection levels. Then the installment at time \( t \) is defined as:

\[
\bar{b}_{t} = b_{t} max\{min\{SI_{x,t}, SI_{max}\}, SI_{min}\}, \tag{2}
\]

where \( b_{t} \) is the basic installment, \( SI_{max} \) and \( SI_{min} \) are, respectively, the maximum and the minimum levels obtained by a reasonable equilibrium between opposite parties’ needs.
Moreover we introduce a profit sharing process, related to the differences between income, capital gains and losses (cf. [3], [8]), so obtaining a balancing goal involving market attractiveness as well as economic sustainability.

This comes true by means of a participation rate applied to the period financial result, when it reaches a certain value, thus generating an embedded option (cf. [4]), by virtue of which, if the period financial result $R_{t+1}$ at the end of the interval $[t, t+1]$, net of the annual quota of the administrative expenses ($\gamma$) is positive, an additional bonus equal to a percentage $\alpha$ of $(R_{t+1} - \gamma)$ is immediately paid to the annuitants or added to the future installments.

For a life annuity with deferment period $T$, premium payment until the time $\tau$, annual installments $\tilde{b}_s$ at the beginning of each year, if the annuitant is alive, assuming that the additional bonus is immediately paid, the total financial result of the $(t+1)$-th accounting period is given by

$$R_{t+1} - \alpha(R_{t+1} - \gamma)^+ = \min(R_{t+1}, (1 - \alpha)R_{t+1} + \alpha\gamma) \quad (3)$$

where

$$R_{t+1} = (V_t + P_{t+1}1_{(t+1 < \tau)})^{-1}(t, t+1) - (\tilde{b}_{t+1} + V_{t+1})1_{(K(x)>t+1)}. \quad (4)$$

To the aim of deeping a profit analysis survey for the proposed life product, it is important to observe that annuities are contracts producing assets generally large if compared with the other life insurance contract typologies. As some Authors suggest (see [6]), in such cases suitable profit measures are indexes as Return on Asset (Roa), the ratio of profit to the assets the business provides, preferably within a policy year. The information concerns the efficiency in using the assets arising from the specific line of insurance products on which we are focusing on, particularly meaningful if the ratio is used with product research purposes.

Schematically Roa index is the ratio between Net Income and Total Asset: thinking about the crucial role the denominator plays in the significance of the index, we will measure the relative efficiency of the participating annuity by means of the amount of profit made per unit of retrospective gain they produce, taking into account the stochastic asset valuation in the form of net retrospective gains arising from the contract. In this acceptation the index can be better indicated as an Actuarial Return on Gain (ARG).

If each contract produces the net cash flow $X_s$ obtained as the difference between premiums and instalments at time $s$ with $s < t$, we can write the expression for the contract retrospective gain valued at time $t$ we will calculate
at the denominator of ARG:

\[ RG_t = \sum_{s=0}^{K_x \wedge t} X_s e^{\int_s^t \delta(u)du} \]

where \( K_x \) is the curtate future lifetime of the \( i \)-th insured and \( x \wedge y = \min(x, y) \). The surplus at time \( t \) is:

\[ S_t = \sum_{j=0}^{\infty} X_j v^{\text{sign}(j-t)(t,j)} \]  

(5)

with

\[ X_j = (P_j + 1)_{(j \leq K_x < T)} - \tilde{b}_j \text{sign}(t - j) \]  

(6)

### 3 Numerical evidences

In what follows we refer to a portfolio of 1000 immediate annuities issued to 1000 lives aged 65. Each insured pays a single premium at the issue time. The pure premium paid by each insured is calculated at a technical rate of 2%.

On the basis of the contractual architecture described in section 2, the assumption is that the benefits due to each insured, if alive at the payment time \( t \), are adjusted by the scale factor defined in (1), where for \( x_p^{\text{tech}} \) we refer to the mortality table SIM2006, whilst we consider the more projected table IPS55 for the \( x_p^{\text{proj}} \) ‘s.

As in [5], in order to limit the impact of the index on the annuity payments, we assume that at most 20% of variation is allowed, than the benefit at time \( t \) is:

\[ \tilde{b}_t = b_t \max\{\min\{S1_{x,t}, 1.2\}, 0.8\} \]

and for the sake of simplicity we set \( b_t = 1 \).

The ARG assessment requires the choice of a stochastic model to describe both the evolution in time of the rate of return obtained by the Company investing the collected premiums and the dynamics of the discounting rate required for quantifying the reserves at each time of valuation.

We choose to model both rates by a Vasicek mean reversion model.

As well known the basic assumption is that the instantaneous spot rate follows the process with constant coefficients under the statistical measure used for historical estimation described by the SDE:

\[ dr_t = \alpha(\mu - r_t)dt + \sigma dW_t \]

with \( \alpha, \mu \) and \( \sigma \) positive constants and \( dW_t \) a standard Wiener process.

Under the risk neutral measure used for valuation and pricing we assume a similar parametrisation for the same process but with one more parameter
in the drift modelling, the market price of risk (cf. [2]). To calibrate the 
Vasicek model we refer to the monthly yield over the period January 2002-
January 2012, on a basket of Treasury Italian bonds listed on the electronic 
bond and government market and having a residual time greater than one 
year.

3.1 ARG assessment in a simulation framework

Aim of this section is the analysis of the stochastic variable ARG as defined 
in section 2, studied in function of $t$. We resort to a simulation procedure 
for getting the empirical distribution of ARG values for each $t$, implemented 
obtaining 10000 ARG outputs for each value of $t$. The ARG expected values 
are attained and, in particular, within a managerial perspective a quantile 
analysis is performed. The quantiles $q(0.99)$ and $q(0.95)$ are shown in Tables 
1 and 2 in order to assess a sort of ”break-even” scenario (depending on the 
investment policy, the choice of the participating level and in general all the 
strategic variables) which presents the ”worst” ARG acceptable in compliance 
with insurer’s opinion expressed by means of a certain confidence level. Figure 
1 shows, in the first subplot, the maximum and minimum values obtained 
for each $t$ by the simulation procedure and the expected value calculated as 
a mean of the 10000 simulated values. In the second subplot of figure 1, the 
quantile values for each $t$ and for each confidence level considered are shown 
together with the expected values. The results refer to a participation quota 
$\alpha = 20\%$.

The clear increasing behavior of the ARG index with time synthetically 
expresses the economic character of the product management and gives the 
measure of how the profit yields the financial resources invested in the product 
itself. The trend presents a steep slope, more marked in the first half of the 
contract duration. A rough collapse of the ARG values is evident when the 
participating quota moves from 20% to 80% (see Tables 1 and 2), being 
understood the increasing trend of the values.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$E[\text{ARG}]$%</th>
<th>$q(0.95)$</th>
<th>$q(0.99)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8927</td>
<td>0.8889</td>
<td>0.8820</td>
</tr>
<tr>
<td>5</td>
<td>0.9886</td>
<td>0.9835</td>
<td>0.9745</td>
</tr>
<tr>
<td>10</td>
<td>1.1093</td>
<td>1.1007</td>
<td>1.0862</td>
</tr>
<tr>
<td>20</td>
<td>1.2819</td>
<td>1.2631</td>
<td>1.2332</td>
</tr>
<tr>
<td>30</td>
<td>1.3676</td>
<td>1.3416</td>
<td>1.3007</td>
</tr>
<tr>
<td>40</td>
<td>1.4083</td>
<td>1.3786</td>
<td>1.3332</td>
</tr>
<tr>
<td>45</td>
<td>1.4083</td>
<td>1.3786</td>
<td>1.3332</td>
</tr>
</tbody>
</table>
Participating variable annuities

Fig. 1. Simulation procedure results: $\alpha = 20\%$

Table 2. Expected ARG and quantiles $\alpha = 0.8$

<table>
<thead>
<tr>
<th>$t$</th>
<th>$E[\text{ARG}]$</th>
<th>$q(0.95)$</th>
<th>$q(0.99)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2442</td>
<td>0.2434</td>
<td>0.2420</td>
</tr>
<tr>
<td>5</td>
<td>0.2627</td>
<td>0.2609</td>
<td>0.2577</td>
</tr>
<tr>
<td>10</td>
<td>0.2829</td>
<td>0.2798</td>
<td>0.2745</td>
</tr>
<tr>
<td>20</td>
<td>0.3102</td>
<td>0.3049</td>
<td>0.2963</td>
</tr>
<tr>
<td>30</td>
<td>0.3229</td>
<td>0.3163</td>
<td>0.3063</td>
</tr>
<tr>
<td>40</td>
<td>0.3287</td>
<td>0.3213</td>
<td>0.3104</td>
</tr>
<tr>
<td>45</td>
<td>0.3295</td>
<td>0.3222</td>
<td>0.3107</td>
</tr>
</tbody>
</table>

References

8  Di Lorenzo et al.

Survival betterment as competitive leverage in insurance sector: profitability analysis for a class of participating variable annuities

Emilia Di Lorenzo\textsuperscript{1}, Albina Orlando\textsuperscript{2}, and Marilena Sibillo\textsuperscript{3}

\textsuperscript{1} Department of Mathematics and Statistics, Via Cinthia Monte S. Angelo, University of Napoli “Federico II”, 80126, Napoli, Italy (E-mail: diloremi@unina.it)
\textsuperscript{2} CNR, Istituto per le Applicazioni del Calcolo Mauro Picone, via P. Castellino 111, 80126 Napoli, Italy (E-mail: a.orlando@iac.cnr.it)
\textsuperscript{3} Department of Economics and Statistics, Campus Universitario, University of Salerno 84084, Fisciano (SA), Italy (E-mail: msibillo@unisa.it)

Abstract. Population aging keeps on creating growth opportunities for life insurance sector; in light of this opportunity, the topics discussed within the paper are focusing on the definition of contractual profiles involving life annuities in order to obtain a reasonable risk-profit sharing between the parties. In particular survival indexed participating annuities are studied, for which a profitability analysis is provided by means of suitable balance-sheet indexes.

JEL Codes: C53, G17, G22, G32
Keywords: Actuarial Return on Gain, variable annuities, longevity risk.

1 Demographic evolution outline and the actuarial products supply

The phenomenon of demographic change develops through several aspects, such as birth decrease and people’s aging, whose management has now become a topic of interest for the international community and constitutes a crucial challenge for today’s society, even with respect to the topic of solidarity among generations. The topic of macro-economic impact of the evolutinal phenomenon of survival has already been brought to the attention of the First World Assembly on Aging, which took place in Wien in 1982; subsequently, the debate kept going on in Madrid in 2002, during the Second World Assembly on Aging (see [9]); today, the issue is constantly being monitored by EU bodies, aiming at transforming a challenge into an opportunity (see [7]).

In particular, among the plethora of questions related to the debated topic, the consequences of population aging within the sectors of social security and public spending are being examined (see: [11]; [7]).
It is appropriate to remind that the European Commission, in its Third Demography Report, published in April 2011, managed to synthesize the growing increase of aging on the basis of the following data: the percentage of the European population older than 65 has increased to 13.7% in 1990; to 17.4% in 2010; projections made for 2060 even estimate a percentage equal to 30%. Furthermore, the so-called "oldest old", namely those individuals who are older than 80, are meant to significantly increase by 2060. Systematic aging will affect the whole Europe, however some regions will particularly be interested, as explained by the studies operated by the Committee of the Regions, recording a variability range which will, in 2030, sway between 10.4% and 37.3% (see [1]). There is little doubt about the magnitude of the impact of this set of data on a socio-economic field.

For what concerns the topics discussed in the paper, it is undoubtedly interesting to focus particularly on the consequences in the insurance sector, from two different perspectives: first of all, in terms of food for thought concerning the actual possibilities of such an important sector and, secondly, in relation to the social relevance of instruments capable of improving life conditions of people of an advanced age.

For what concerns the current size of insurance sectors, instead, Swiss Re provided a precise mapping of insurance industry on a global scale, which allows for a clear understanding of the consequences of the aging phenomenon in relation to the demand for aimed products (see [10]). The aforementioned report, based upon the data collected in 78 countries, representing 98% of global premium collection, highlights the positive trend of insurance sector, regardless of the unavoidable criticalities related to investment returns; in fact, the volume of premiums has globally registered an increase by 2.7%, thus resulting into a sum, which, despite being inferior to those preceding the financial crisis, exceeds 4300 billion dollars. For what concerns life sector, instead, Swiss Re’s report shows how population aging keeps on and will keep on creating growth opportunities; this represents, then, a set of opportunities which are to be strongly defended if we think about the current financial condition and, mainly, about public debt crisis in Europe - extremely dangerous for actives - and the complex regulatory implementation of Solvency II. On the basis of what explained, then, it seems evident that among the many peculiarities characterizing the complex demographic change phenomenon, the evolitional dynamics of survival, characterized, during the last decades, as well known, by a certainly positive trend, do occupy an important position, due to the consequences deriving from it in an economic, social and cultural environment. The topics discussed within the paper are mainly focusing on a sector which is closely connected to the matter concerning the evolution of survival, namely the security-insurance one. In this sense, the aging phenomenon has been studied as an opportunity for development of a proper financial sector, with the positive consequence of a competitive leverage. Such line of inquiry is not, anyway, unrelated to social wellness.
improvement goals, since it concretely responds to the needs of a relevant segment of the population, currently under the spotlight at a European level. In fact, the words stated by the European platform Age (see [1]), concerning senior consumers rights, are particularly meaningful: "to guarantee older people’s rights as consumers as prescribed in the EU legal framework and to help raise awareness regarding those rights and how to make better decisions as consumers”.

Therefore, on the basis of what pointed out earlier, the impact of innovative contractual architectures aimed at risk-profit sharing between the parties, framed within the picture represented by the scenario described earlier, has a multi-faceted relevance; in detail, they may provide useful instruments, which would bring advantage to economic growth, on one hand on the level of insurance industry development (which now finds itself forced to face a growing demand of contractual typologies, appropriate to evolutional dynamics of survival; in particular, one goal stands above the others: giving an appropriate response to the increasing pension sector demands), on the other hand, on the level of safeguard for those who ask for investment instruments and possibilities fitting an elderly age. The afore mentioned advantages have to be framed within the development of competitive knowledge in a financial and actuarial field, where such a development undoubtedly implies a conscious use of the instruments proposed by financial and insurance engineering, which is largely fuelled by correct knowledge and scientific innovation.

2 Contractual architecture within a risk-profit sharing perspective

In the following we consider profit participating life annuities with installments scaled by a demographic index structured on the forecasted survival trend and the observed one. According to the contractual scheme introduced by Denuit et al. ([5]), the installments for to each insured aged $x$ at issue, if alive at the payment time $t$, are scaled by the factor:

$$SI_{x,t} = \frac{\xi P_t^\text{tech}}{\xi P_t^\text{proj}},$$

where $\xi P_t^\text{tech}$ represents the survival probability for an insured aged $x$ to age $x+t$ inferable from a proper technical bases and $\xi P_t^\text{proj}$ the survival probability inferable from a projected demographic model.

Following Denuit et al. ([5]), we also introduce a minimum and a maximum threshold for the survival index, so overcoming marketing problems arising from increasing projection levels. Then the installment at time $t$ is defined as:

$$\tilde{b}_t = b_t \max\{\min\{SI_{x,t}, SI_{\text{max}}\}, SI_{\text{min}}\},$$

where $\xi P_t^\text{tech}$ represents the survival probability for an insured aged $x$ to age $x+t$ inferable from a proper technical bases and $\xi P_t^\text{proj}$ the survival probability inferable from a projected demographic model.
where \( b_t \) is the basic installment, \( SI_{\text{max}} \) and \( SI_{\text{min}} \) are, respectively, the maximum and the minimum levels obtained by a reasonable equilibrium between opposite parties’ needs.

Moreover we introduce a profit sharing process, related to the differences between income, capital gains and losses (cf. [3], [8]), so obtaining a balancing goal involving market attractiveness as well as economic sustainability.

This comes true by means of a participation rate applied to the period financial result, when it reaches a certain value, thus generating an embedded option (cf. [4]), by virtue of which, if the period financial result \( R_{t+1} \) at the end of the interval \([t, t+1]\), net of the annual quota of the administrative expenses \( (\gamma) \) is positive, an additional bonus equal to a percentage \( \alpha \) of \( (R_{t+1} - \gamma) \) is immediately paid to the annuitants or added to the future installments.

According to an ex-ante analysis at the issue time, information about interest rates and survivors’ number at time \( t \) are incorporated within \( F = (F(s))_{s \in \{0,1,\ldots,t\}} \) and \( M = (M(s))_{s \in \{0,1,\ldots,t\}} \), respectively, so generating the filtration \( H = (H(s))_{s \in \{0,1,\ldots,t\}} \), with \( H(s) = F(s) \cup M(s) \).

For a life annuity with deferment period \( T \), premium payment until the time \( \tau \), annual installments \( \tilde{b}_s \) at the beginning of each year, if the annuitant is alive, assuming that the additional bonus is immediately paid, the total financial result of the \((t+1)\)-th accounting period is given by

\[
R_{t+1} - \alpha(R_{t+1} - \gamma)^+ = \min(R_{t+1}, (1-\alpha)R_{t+1} + \alpha\gamma) \tag{3}
\]

where

\[
R_{t+1} = (V_t + P_{t+1}1_{(t+1<\tau)})v^{-1}(t, t+1) - (\tilde{b}_{t+1} + V_{t+1})1_{(K(x)>t+1)}. \tag{4}
\]

To the aim of deeping a profit analysis survey for the proposed life product, it is important to observe that annuities are contracts producing assets generally large if compared with the other life insurance contract typologies. As some Authors suggest (see [6]), in such cases suitable profit measures are indexes as Return on Asset (Roa), the ratio of profit to the assets the business provides, preferably within a policy year. The information concerns the efficiency in using the assets arising from the specific line of insurance products on which we are focusing on, particularly meaningful if the ratio is used with product research purposes.

Schematically Roa index is the ratio between Net Income and Total Asset: thinking about the crucial role the denominator plays in the significance of the index, we will measure the relative efficiency of the participating annuity by means of the amount of profit made per unit of retrospective gain they produce, taking into account the stochastic asset valuation in the form of net retrospective gains arising from the contract. In this acceptation the index can be better indicated as an Actuarial Return on Gain (ARG).

If each contract produces the net cash flow \( X_s \) obtained as the difference between premiums and instalments at time \( s \) with \( s < t \), we can write the ex-
pression for the contract retrospective gain valued at time $t$ we will calculate at the denominator of $\text{ARG}$:

$$\text{RG}_t = \sum_{s=0}^{K_{x_i} \wedge t} X_s e^{\int_{s}^{t} \delta(u)du}$$

where $K_{x_i}$ is the curtate future lifetime of the $i$-th insured and $x \wedge y = \min(x, y)$. The surplus at time $t$ is:

$$S_t = \sum_{j=0}^{\infty} X_j v^{\text{sign}(j-t)(t,j)}$$

with

$$X_j = (P_{j+1} 1_{j \leq K_x < T} - \tilde{b}_j) \text{sign}(t-j)$$

### 3 Numerical evidences

In what follows we refer to a portfolio of 1000 immediate annuities issued to 1000 lives aged 65. Each insured pays a single premium at the issue time. The pure premium paid by each insured is calculated at a technical rate of 2%.

On the basis of the contractual architecture described in section 2, the assumption is that the benefits due to each insured, if alive at the payment time $t$, are adjusted by the scale factor defined in (1), where for $x_{p\text{tech}}^t$ we refer to the mortality table SIM2006, whilst we consider the more projected table IPS55 for the $x_{p\text{proj}}^t$’s.

As in [5], in order to limit the impact of the index on the annuity payments, we assume that at most 20% of variation is allowed, than the benefit at time $t$ is:

$$\tilde{b}_t = b_{t_{\max}} \{\min\{S_{I_{x,t}}, 1.2\}, 0.8\}$$

and for the sake of simplicity we set $b_t = 1$.

The ARG assessment requires the choice of a stochastic model to describe both the evolution in time of the rate of return obtained by the Company investing the collected premiums and the dynamics of the discounting rate required for quantifying the reserves at each time of valuation.

We choose to model both rates by a Vasicek mean reversion model.

As well known the basic assumption is that the instantaneous spot rate follows the process with constant coefficients under the statistical measure used for historical estimation described by the SDE:

$$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t$$

with $\alpha$, $\mu$ and $\sigma$ positive constants and $dW_t$ a standard Wiener process.
Under the risk neutral measure used for valuation and pricing we assume a similar parametrisation for the same process but with one more parameter in the drift modelling, the market price of risk (cf. [2]). To calibrate the Vasicek model we refer to the monthly yield over the period January 2002-January 2012, on a basket of Treasury Italian bonds listed on the electronic bond and government market and having a residual time greater than one year.

3.1 ARG assessment in a simulation framework

Aim of this section is the analysis of the stochastic variable ARG as defined in section 2, studied in function of $t$. We resort to a simulation procedure for getting the empirical distribution of ARG values for each $t$, implemented obtaining 10000 ARG outputs for each value of $t$. The ARG expected values are attained and, in particular, within a managerial perspective a quantile analysis is performed. The quantiles $q(0.99)$ and $q(0.95)$ are shown in Tables 1 and 2 in order to assess a sort of ”break-even” scenario (depending on the investment policy, the choice of the participating level and in general all the strategic variables) which presents the ”worst” ARG acceptable in compliance with insurer’s opinion expressed by means of a certain confidence level. Figure 1 shows, in the first subplot, the maximum and minimum values obtained for each $t$ by the simulation procedure and the expected value calculated as a mean of the 10000 simulated values. In the second subplot of figure 1, the quantile values for each $t$ and for each confidence level considered are shown together with the expected values. The results refer to a participation quota $\alpha = 20\%$ and $\alpha = 80\%$.

The clear increasing behavior of the ARG index with time synthetically expresses the economic character of the product management and gives the measure of how the profit yields the financial resources invested in the product itself. The trend presents a steep slope, more marked in the first half of the contract duration. A rough collapse of the ARG values is evident when the participating quota moves from 20% to 80% (see Tables 1 and 2), being understood the increasing trend of the values.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$E[\text{ARG}]%$</th>
<th>$q(0.95)$</th>
<th>$q(0.99)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8927</td>
<td>0.8889</td>
<td>0.8820</td>
</tr>
<tr>
<td>5</td>
<td>0.9886</td>
<td>0.9835</td>
<td>0.9745</td>
</tr>
<tr>
<td>10</td>
<td>1.1093</td>
<td>1.1007</td>
<td>1.0862</td>
</tr>
<tr>
<td>20</td>
<td>1.2819</td>
<td>1.2631</td>
<td>1.2332</td>
</tr>
<tr>
<td>30</td>
<td>1.3676</td>
<td>1.3416</td>
<td>1.3007</td>
</tr>
<tr>
<td>40</td>
<td>1.4083</td>
<td>1.3786</td>
<td>1.3332</td>
</tr>
<tr>
<td>45</td>
<td>1.4083</td>
<td>1.3786</td>
<td>1.3332</td>
</tr>
</tbody>
</table>
Participating variable annuities 7

Fig. 1. Simulation procedure results: $\alpha = 20\%$

Table 2. Expected ARG and quantiles $\alpha = 0.8$

<table>
<thead>
<tr>
<th>$t$</th>
<th>$E[\text{ARG}]%$</th>
<th>$q(0.95)$</th>
<th>$q(0.99)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2442</td>
<td>0.2434</td>
<td>0.2420</td>
</tr>
<tr>
<td>5</td>
<td>0.2627</td>
<td>0.2609</td>
<td>0.2577</td>
</tr>
<tr>
<td>10</td>
<td>0.2829</td>
<td>0.2798</td>
<td>0.2745</td>
</tr>
<tr>
<td>20</td>
<td>0.3102</td>
<td>0.3049</td>
<td>0.2963</td>
</tr>
<tr>
<td>30</td>
<td>0.3229</td>
<td>0.3163</td>
<td>0.3063</td>
</tr>
<tr>
<td>40</td>
<td>0.3287</td>
<td>0.3213</td>
<td>0.3104</td>
</tr>
<tr>
<td>45</td>
<td>0.3295</td>
<td>0.3222</td>
<td>0.3107</td>
</tr>
</tbody>
</table>

References

1. AGE Platform Europe in partnership with the Committee of the Regions and the European Commission. How to promote active ageing in Europe. EU support to local and regional actors (2011).


4. V. D’Amato, E. Di Lorenzo, A. Orlando, M. Russolillo and M. Sibillo. Profit Participation Annuities: A Business Profitability Analysis within a Demographic
8  Di Lorenzo et al.


Farm Irrigation Water Technical Efficiency: Stochastic Frontier Analysis

Elgilany A. Ahmed, Jamalludin Sulaiman, Saidatulakmal Mohd, Okezie, A. Chukwukere
Department of Economics, School of Social Science, University Sains Malaysia, School of Social Sciences, Pulau Pinang, Malaysia

E-mail: elgilanya@yahoo.com or algaylani@usm.my

Abstract
Pump-irrigated food crops are grown in the River Nile State (RNS), which offers a suitable environment for its production. Cereal crops such as wheat, commonly grown in these schemes faces a manifold problem, namely inefficient irrigation-water use coupled with low productivity level and high production costs. The study investigated On-Farm Water-Use technical efficiency for Wheat production. Systematic sampling was adopted in selecting 70 scheme participants and data elicited through a cost route approach. Cropwat4 analysis and stochastic frontier model were use in analyzing water use and efficiency in Wheat production. Results revealed that the FWUE in wheat production was only 64%, while growers were over-irrigating the crop by 36%. High irrigation-water cost is the most critical production constraint, which accounted for 19% of wheat’s total variable production costs. There was inefficiency in water use and the significant determinants of output of Wheat are water price and fertilizer. The inefficiency factors include farming experience, extension visits, irrigation distance and off farm income. Thus farmers were inefficient in the use of production factors as they operated below the frontier output. There is enormous potential for improvements that could lead to substantial savings in irrigation water, which in turn can be utilized to gain additional irrigated areas.

Keywords: Water-use, Efficiency, Production costs, water productivity, Wheat

Introduction
Irrigation systems for agriculture, viticulture, and residential and commercial landscapes consume enormous amounts of water resources. The convenience of automatic irrigation systems often results in over-watering and associated negative impacts of mineral and fertilizer leaching and above ground runoff. An effective irrigation system needs to apply the appropriate amount of water without applying too much (Fazackerley et al, 2011). Wheat (*Triticum aestivum*) is a strategic crop in Sudan. It is Sudan's second most important cereal food in terms of consumption after sorghum. Over the past few years, wheat production, which is almost entirely irrigated, has been declining due to diminishing yields and soaring input costs. Since 1999, the Government liberalized the wheat production regime and removed all support programs. These moves have prompted many farmers to drastically reduce wheat cultivation and/or switch to more lucrative cash crops, such as vegetables and oil seeds. Gezira, White Nile, New Halfa, River Nile and Northern States are the main suppliers for irrigated wheat. The overall area under wheat in 2005/06 exceeded 290 thousand feddans (122 000) hectares (MAS, 2006). The last three decades witnessed critical problems regarding water provision, distribution and utilization particularly encountered at peak demand periods that may be attributed to power failure accompanied by lack and high cost of fuel and spare parts to operate the pumps. These became common problems and were aggravated even more by the diminution of the canals.
carrying capacities, resulting in low productivity of crops. Faki (2004) stated that the high cost of production, coupled with low productivity and lack of a cheap source of power, has made it difficult for the farmers to realize the full potential of the State. Further, development is constrained by serious limitations on the two basic resources: land and water. Flow recession of the Nile and Atbara Rivers in August and September affect the availability of irrigation water in October through February. Requirement for irrigation water is highest through this period when winter cash crops (i.e. wheat, legumes, fodders, vegetables) are grown and require irrigation water. According to Elsir et al. (2004), it has long been recognised that high production costs, low productivity and lack of a cheap source of power for water pumping hinder realisation of the full agricultural potential in River Nile State. The awareness of water use efficiency not widespread in most of the public schemes in the State; in other words there is low strive to apply the recommended standards of crop water requirements. The area that can be commanded by pumps is significantly higher than the actually cultivated one. This indicates that the capacity of those pumps was underutilized. Faki (2004) reported that in Northern Sudan the irrigation needs are designated in terms of numbers of irrigations, not actual quantities, and it is likely that reduction in amounts per irrigation or even number of irrigations may be possible without reducing yield. Surface irrigation by pumps, which dominates the RNS agricultural production systems, is regarded as having very low efficiency, leading to low FWUE in the majority of the public irrigated schemes there. Therefore, this paper examines water use and productivity of Wheat in the public irrigation schemes of the River Nile State in Sudan.

Methodology
Study Area: The River Nile State is considered as one of the poorest States in the country, although, with its relatively cooler weather and fertile alluvial soils, has a comparative advantages over other parts of the country in producing relatively high-value crops (wheat, faba beans, citruses, mangos, dates, certain spices and medical plants). Wheat is regarded as one of the major winter food crops in the State; the total area planted under wheat in season 2005/06 is estimated at 75,434 feddans. The State accommodates three types of pump-irrigated schemes: private, cooperative and public schemes with different production-relation systems. The optimum time for wheat planting is November and temperatures during this period remained cooler than average. In almost all wheat growing areas, the bulk of plantings occur on time. The effectiveness of irrigation is determined by the availability and supply of fuel for pump irrigation and the degree of siltation in canals for larger schemes.

Sampling and Data Collection
Stratified random sampling was employed in selecting 70 respondents for the study, forming about 2.3% of Elzeidab scheme’s tenants. The cost route approach was adopted in collecting data for the 2005/2006 season. Data collection involved personal interviews and the use of structured questionnaire.

Data Analysis
Data analysis involved descriptive techniques and the use of Crop Wat 4. Water use efficiency was analysed by the stochastic frontier model.

Economics of Wheat Production
Gross Margin Analysis (GMA) was employed to determine the economics of Wheat production. The gross, which is calculated as gross revenue less variable costs forms a
good indicator of how profitable a firm is at the most fundamental level. The general mathematical form for the gross margin analysis used to calculate the gross margin as follow:

\[ GM = GR - TVC \] 

Where:
- \( GM \): gross margin of each crop per feddan in SD.
- \( GR \): gross revenue of each crop per feddan in SD.
- \( TVC \): total variable cost per feddan in SD.

Assessment of irrigation water use efficiency for wheat
Assessment of water use under full irrigation provides important indicators for WUE in Wheat production. According to ICARDA (2001) the concept of on-farm water use efficiency (FWUE) was developed to address this complex situation at the farm level. FWUE is defined as the ratio of the required irrigation water to produce a specific output level to the actual amount of water applied by farmers. With this definition, FWUE may take the value of less, equal or greater than one. A value less than one implies that farmers over-irrigate their crops, while a value greater than one implies that farmers under-irrigate their crops. However, if the value of the calculated FWUE is equal to one, it means that farmers are fully efficient in using irrigation water because the required and applied amounts of water are equal, as shown in the following form:

\[ FWUE = \frac{W_r}{W_a} \times 100 \]

Where:
- \( W_r \): is the amount of water required (m³) by the crop to produce a certain level of crop production.
- \( W_a \): is the amount of water actually applied (m³) by farmers to produce that level of crop production.

Stochastic frontier production model specification
The Stochastic Frontier Analysis (SFA) investigates farm specific determinants of productivity. The model computes efficiency values as indicators of productivity and determinants of efficiency. This approach was adopted to determine the effect of irrigation water use by farmers on the output of Wheat.

A general stochastic production frontier model can be given by:

\[ \ln q_j = f(\ln x) + v_j - u_j \]

Where \( q_j \) is the output produced by firm \( j \), \( x \) is a vector of factor inputs, \( v_j \) is the stochastic (white noise) error term and \( u_j \) is a one-sided error representing the technical inefficiency of firm \( j \). Both \( v_j \) and \( u_j \) are assumed to be independently and identically distributed (iid) with variance \( \sigma_v^2 \) and \( \sigma_u^2 \) respectively.

Given that the production of each firm \( j \) can be estimated as:
\[ \ln \hat{q}_j = f(\ln x) - u_j \]  \[ \text{4} \]

While the efficient level of production (i.e. no inefficiency) is defined as:

\[ \ln q^* = f(\ln x) \]  \[ \text{5} \]

Then technical efficiency (TE) can be given by:

\[ \ln TE_j = \ln \hat{q}_j - \ln q^* = -u_j \]  \[ \text{6} \]

Hence,

\[ TE_j = e^{-u_j} \]  \[ \text{7} \]

-Equation (7) - and is constrained to be between zero and one in value. If \( u_j \) equals zero, then TE equals one, and production is said to be technically efficient. Technical efficiency of the \( j \)th firm is therefore a relative measure of its output as a proportion of the corresponding frontier output. A firm is technically efficient if its output level is on the frontier, which implies that \( \frac{q}{q^*} \) equals one in value.

One aim in SFA is to explain inefficiency/efficiency in terms of exogenous determinants, (Kumbhakar and Lovell, 2000) summarized some models to explain inefficiency/efficiency of a producer. The generalized efficiency equation is given as:

\[ U_i = g(zi; \gamma) + \epsilon_i \]  \[ \text{8} \]

Results and Discussion

Socio-demographic characteristics of Respondents: The average age of surveyed tenants in Elzeidab scheme was 40 years, while the average family size of surveyed tenants ranged from 1 to 15 persons. The study detected that all surveyed tenants were educated, and all of them were males. Farming experience estimated at 20 years on the average. The average farm size in the scheme varied from 1 to 28 feddan per farm household. The average distance from tenant’s residence to their farms was 2.7 km. The farming system of Elzeidab scheme is dominated by wheat production which accounts for 25% of the farm land.

Elzeidab irrigation-water costs

According to MAS report (2006) irrigation water costs for Elzeidab scheme for season 2005/2006 are broadly differentiated into fixed and variable costs (see Figure 1). Irrigation costs are mainly based on variable costs, while MAS often ignores the fixed costs. The annual running expenses comprise fuel (diesel), oil (engine, gearbox), grease, spare parts, maintenance, staff salaries and allowances (management expenses), services and others. The analysis revealed that the fuels item amounted to SD 116,422,848 forming the highest item of the cost of irrigation reaching 53% of the total irrigation costs.
Main irrigation variable cost items for Elzeidab scheme

Figure 1: the irrigation cost items in percentage

Irrigation-water cost is invariably considered as the most critical agricultural constraint on account of the high cost of water pumping from the RN. Detailed illustration of disaggregated production costs is shown in Figure 2 and depicts the cost items in sequence of the seasonal crop production operations. The 13 cost components shown in Figure 2 add up to SD 70054 as the total cost of production. The irrigation-cost component formed the highest cost item accounting for 19% of total production cost. Wheat growers in the scheme pay the cost of this item as a fixed rate to the scheme administration at the end of the season.

Figure 2: Percentage shares of wheat variable cost items

Gross Margin Analysis
According to the survey results, wheat production costs were less than its gross returns resulting in a positive gross margin of SD 4295 in season 2005/06 (Table 1).

Table 1: Gross margin analysis of wheat in Elzeidab scheme

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production cost (SD/feddan)</td>
<td>70054</td>
</tr>
<tr>
<td>Average yield (kg/feddan.)</td>
<td>676</td>
</tr>
<tr>
<td>Average price (SD/kg)</td>
<td>110</td>
</tr>
<tr>
<td>Gross return (SD/feddan.)</td>
<td>74349</td>
</tr>
<tr>
<td>Gross marginal revenue (SD/feddan.)</td>
<td>4295</td>
</tr>
</tbody>
</table>
The Table also shows that, although the gross margin of wheat was positive, it was nevertheless low, especially if the forgone opportunities of using winter land and water are considered, given the wide range of winter crops that can be produced. One reason for the low gross margin is the increasing input prices faced in the RNS in the last decade. According to this fact, wheat could be assessed as infeasible crop unless improvements are made.

**Irrigation water use efficiency**
FWUE of wheat was estimated at two levels: per watering and per season. The average water application per season for average wheat was found to be 3756 m$^3$/feddan as depicted in Table 2.

<table>
<thead>
<tr>
<th>Crop</th>
<th>CWa (m$^3$/fed)</th>
<th>FWUE - watering</th>
<th>Over-irrigation %</th>
<th>FWUE -season</th>
<th>Over-irrigation %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td>3756</td>
<td>0.41</td>
<td>59</td>
<td>0.64</td>
<td>36</td>
</tr>
</tbody>
</table>

The analysis revealed that FWUE for wheat is not high if compared to some seasonal crops in the scheme which consume relatively high amounts. Onions, vegetables and potatoes are high water-demanding crops throughout their growing seasons that extend over about 141, 130 and 110 days, respectively. The estimated FWUE of Elzeidab scheme indicated a wide technological gab between the required utilization and actual water application for wheat and other crops. The Figure also shows that farmers within the surveyed sample over-irrigated their entire field crops.

The average cultivated farm area in the area of study, which was 6.0149 feddan per tenant, while the average area of cultivated wheat was 3.672 feddan, or 61% of the total cultivated farm area. Wheat growers in the scheme exceeded the crop water requirements by 41% per watering and 64% for the whole season, suggesting the need for improving the FWUE. Further, the study quantified the amount of water supplied by the scheme and the quantity required for Wheat production as depicted in Table 3.

<table>
<thead>
<tr>
<th>category</th>
<th>Av. cultivated area (fed)</th>
<th>CWa (m$^3$)</th>
<th>CWr (m$^3$)</th>
<th>irrigation gap (m$^3$)</th>
<th>Expected extension area (fed)</th>
<th>Expected extension area (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farm size 8.509 fed.</td>
<td>6.0149</td>
<td>28573</td>
<td>13432</td>
<td>15140.59</td>
<td>6.779</td>
<td>112</td>
</tr>
<tr>
<td>Wheat</td>
<td>3.672</td>
<td>13792</td>
<td>8800</td>
<td>4991.72</td>
<td>2.083</td>
<td>57</td>
</tr>
</tbody>
</table>

The average amount of water available to the total cultivated farm area was 28573 m$^3$; while the average quantity of crop water requirements was 13432 m$^3$, indication surplus water of 15141 m$^3$, which is sufficient for supplying expected extensions in irrigated areas derived at 6.779 fed or 112% of the farm cultivated area. Table 3 also shows the average
amount of water available to wheat within the farm amounting to 13792 m$^3$. The average quantities of crop water requirements were 8800 m$^3$ with an estimated surplus water of 4992 m$^3$; sufficient for a possible extension in irrigated area of 2.083 fed equal equivalent to 57% of cultivated area.

**Productivity per unit water for wheat**

Cropping pattern is one of the most important parameters involved in irrigation command areas. It is directly related to the productivity of irrigation systems and greatly contributes to improved soil and water utilization. Crop planning in irrigated agriculture has traditionally been based on the concept of maximization of net benefit (Montazar *et al.*, 2011). According to ICARDA research on WUE, water productivity is defined as the ratio of crop production (kg) to the unit of water used (m$^3$) or as the amount of food produced per unit volume of water used. There are several different ways of expressing water productivity such as pure physical productivity, combined physical and economic productivity. Determination of productivity per unit water of wheat here was assessed for both economic and physical productivity of water. Water productivity in monetary terms (in SD) of output per m$^3$ of water as depicted in Table 4 provides an important indicator of water productivity.

**Table 4: Determination of wheat productivity per unit water in monetary terms for the surveyed tenants in Elzeidab scheme**

<table>
<thead>
<tr>
<th>Crops</th>
<th>Water price (SD)</th>
<th>Amount of water (m$^3$/fed)</th>
<th>Water productivity (SD/m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td>13592.59</td>
<td>3756.00</td>
<td>3.619</td>
</tr>
</tbody>
</table>

From Table 4, water productivity for wheat in monetary term was only 3.619 SD/m$^3$, while it was 5.452 SD/m$^3$ for spices. On the other hand, physical (or technical) water productivity that measures kgs of output per m$^3$ of water was only 0.180 kg/m$^3$, which very low when compared to water productivity for potatoes derived at 0.680 kg/m$^3$ in the area of study (Table 5). Hence water productivity in technical or economic terms have important considerations for the assessment of crops produced in the irrigated subsector.

**Table 5: Determination of wheat productivity per unit water in physical terms for the surveyed tenants in Elzeidab scheme season 2005/06**

<table>
<thead>
<tr>
<th>Crops</th>
<th>Yield (kg)</th>
<th>Amount of water (m$^3$/fed)</th>
<th>Productivity per unit water (kg/m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat in</td>
<td>676</td>
<td>3756</td>
<td>0.180</td>
</tr>
<tr>
<td>the scheme</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Determinants of efficiency in wheat production**

The frontier result of the determination of efficiency is shown in table 6. The model statistics estimated are all valid. The value of gamma (g) indicates the proportion of variation in the model that is due to capacity production factors included in the model. The value is relatively high, 0.75 percent and is statistically significant at one percent. The implication is that most of the variables included in the model are necessary in accounting for the output of Wheat in Sudan.
Table 6: Results from maximum likelihood estimation

<table>
<thead>
<tr>
<th>Production factors</th>
<th>Coefficient</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant ($\beta_0$)</td>
<td>5.397***</td>
<td>8.30</td>
</tr>
<tr>
<td>Land ($\beta_1$)</td>
<td>-0.031</td>
<td>-1.26</td>
</tr>
<tr>
<td>Seed ($\beta_2$)</td>
<td>0.005</td>
<td>0.47</td>
</tr>
<tr>
<td>Water Price ($\beta_3$)</td>
<td>-0.051***</td>
<td>-3.740</td>
</tr>
<tr>
<td>Capital ($\beta_4$)</td>
<td>-0.026</td>
<td>-0.08</td>
</tr>
<tr>
<td>Fertilizer ($\beta_5$)</td>
<td>0.001**</td>
<td>2.21</td>
</tr>
<tr>
<td>Animal Power ($\beta_6$)</td>
<td>0.001</td>
<td>1.51</td>
</tr>
<tr>
<td>Labour cost ($\beta_7$)</td>
<td>0.001</td>
<td>1.12</td>
</tr>
<tr>
<td>Efficiency factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education of household head ($\delta_1$)</td>
<td>0.127</td>
<td>0.81</td>
</tr>
<tr>
<td>Household size ($\delta_2$)</td>
<td>-0.129</td>
<td>-0.50</td>
</tr>
<tr>
<td>Farming experience ($\delta_3$)</td>
<td>0.352*</td>
<td>1.82</td>
</tr>
<tr>
<td>Age of household head ($\delta_4$)</td>
<td>-0.111</td>
<td>-0.29</td>
</tr>
<tr>
<td>Extension Visit/Access ($\delta_5$)</td>
<td>-0.289*</td>
<td>-1.78</td>
</tr>
<tr>
<td>Irrigation distance ($\delta_6$)</td>
<td>0.695**</td>
<td>2.09</td>
</tr>
<tr>
<td>Home to field distance ($\delta_7$)</td>
<td>-0.409</td>
<td>-0.73</td>
</tr>
<tr>
<td>Off-farm Income ($\delta_8$)</td>
<td>-0.624*</td>
<td>-1.95</td>
</tr>
<tr>
<td>sigma-squared</td>
<td>0.49**</td>
<td>2.51</td>
</tr>
<tr>
<td>Gamma (g)</td>
<td>0.75***</td>
<td>5.06</td>
</tr>
<tr>
<td>log likelihood function</td>
<td>-52.69</td>
<td></td>
</tr>
<tr>
<td>LR test of the one-sided error</td>
<td>14.62</td>
<td></td>
</tr>
</tbody>
</table>

The generalized likelihood ratio statistic (also known as the LR test) is high 14.62 which led us to conclude that the production frontier is identical to the production function. The results obtained are valid and not spurious. Water price for irrigation and fertilizer are the significant determinants of output of Wheat. While water price is negative and significant at 1 percent, fertilizer is a positive determinant of output and is significant at 5 percent. Previous studies by Amaza (2000), Adeoti (2001) Ogundele (2003) and Awotide (2004) also reported low elasticity for fertilizer in food crop production in Nigeria. The economic price of irrigation water is not charged and this has resulted in large volumes of water been wasted. Farmers are not charged on the basis of volume used and once they pay subject to use the way they want. Wheat growth depends on high soil fertility and as such there is heavy application of fertilizer for high productivity. Farmers enhance the productivity of land through the application of fertilizers. Four factors were significant in the inefficiency model viz farming experience, Extension visits, irrigation distance and off-farm income. In the inefficiency model, a negative coefficient implies an increase in efficiency while a negative coefficient leads to reduction in efficiency other things being equal. As the farmer gets older he becomes less innovative and takes little risk. This accounts for the impact of farming experience. Extension visits impacts positively on efficiency. Farmers receive advice of modern techniques and advice about the pest way to handle farm problems and availability of modern seeds and seedlings. The further away the source of irrigation, the more difficult farmers make use of the service. Farmers provide their own transport and as such when the source is far away, they spend more money on transportation and use less of the services. This accounts for while distance impacts negatively on efficiency. People who...
earn off farm income and get remittance often do not concentrate in terms of following all the agronomic practices and this often decreases productivity hence efficiency.

**Technical efficiency of wheat farmers**

Irrigation development requires the mobilization of often scarce resources, including arable land, adequate water, and financial capital. Hence, it is imperative that the organization subsequently entrusted with responsibility for the completed irrigation scheme, be capable of ensuring proper management of such resources. Efficient use of these resources may be evaluated using indicators such as: cropping intensity (IC), the proportion of area affected by damage (PSD) and relative water supply (RWS) (Dembele *et al.*, 2011). The frequency distribution of the technical efficiency indices derived from the analysis of the stochastic production is provided in Table 7.

**Table 7: Distribution of economic efficiency**

<table>
<thead>
<tr>
<th>Efficiency Range</th>
<th>Mean</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.14 – 0.30</td>
<td>0.208</td>
<td>7.14</td>
</tr>
<tr>
<td>0.31 – 0.47</td>
<td>0.421</td>
<td>15.71</td>
</tr>
<tr>
<td>0.48 – 0.64</td>
<td>0.582</td>
<td>32.86</td>
</tr>
<tr>
<td>0.65 – 0.81</td>
<td>0.722</td>
<td>31.43</td>
</tr>
<tr>
<td>0.82 – 1.00</td>
<td>0.866</td>
<td>12.86</td>
</tr>
<tr>
<td>Grand Mean</td>
<td>0.611</td>
<td></td>
</tr>
</tbody>
</table>

The technical efficiency of the sampled farmers was less than one (or 100%) indicating that all the maize farmers sampled were operating below the frontier. The best performing farm had a technical efficiency of 0.95 or 95 percent, while the least performing farm had a technical efficiency of 0.14 or 14 percent. The mean technical efficiency of the Wheat farmers was 0.61 or 61 percent. This implied that the Wheat farmers were able to obtain about 61 per cent of optimal output from a given set of production inputs suggesting that there is the scope for increasing Wheat production by 39 per cent if they were to operate at the frontier or by 5 percent if all Wheat farmers would adopt the technology and production techniques currently used by the most technically efficient farmer. In general, the results suggested that the sampled farmers were fairly technically efficient.

**Conclusion**

Wheat has been a very important crop in the Sudan agricultural production system. Most of the Wheat produced is under irrigation system. Water use is critical in the production of Wheat and therefore the study investigated the efficiency of Wheat Water use in the Elzeidab scheme. A systematic procedure was adopted in selecting scheme participants and a cost route approach was adopted in data collection. Analytical procedure in studying Wheat water efficiency and productivity include gross margin analysis for profitability, cropWAT analysis for water use efficiency and the stochastic frontier analysis for the determinants of efficiency in Wheat production. The farming system of Elzeidab scheme is dominated by wheat, which accounts for 25% of the farm land. With a share of 53%, fuel was the highest single cost item of irrigation costs. On the other hand, wheat irrigation costs formed 19% of its total cost of production representing the highest category of overall variable costs. Awareness about CWR is absent among all surveyed tenants; a situation that
would be attributable to limitation in extension services. The estimated FWUE indicated a wide technological gab between the CWR and the actual applied water; reaching 41% per watering and 64% for the entire season. The estimated wasted amounts of irrigation water would be sufficient for expected irrigated-area extension that is determined as 57% of the average grown area of wheat and 112% of the average cultivated area of the scheme. Water productivity of wheat crop in monetary and physical terms was generally low in the area of study when compared to some seasonal crops. Water price for irrigation and fertilizer were the significant determinants of the output of Wheat in the scheme. While the inefficiency factors are farming experience, extension visits, irrigation distance and off-farm income. The average efficiency for the scheme is 61 percent showing that farmers are producing below the frontier output level. Based on obtained results, the noted the high amount of water wastage and recommends that building capacity for huge water savings. This capacity should be utilized for expansions in uncultivated areas in the State through State intervention and adoption of participatory approaches involving scheme administrators and tenants to manage irrigation water and sensitizing to adopt modern water saving technologies.

References


MAS, 2006. Annual report prepared by the State Ministry of Agriculture and Irrigation.


Web-Sampling: Using The Moving Average For Data Presentation

Nikolaos Farmakis and George Makris
Aristotle University of Thessaloniki, Dept of Mathematics, Webscience Program, Greece
Email: farmakis@math.auth.gr
Aristotle University of Thessaloniki, Dept of Mathematics, Webscience Program, Greece
Email: geormak@hotmail.com

Abstract: The Internet is the basic tool of the international informatics society. Trillions of information pieces are exchanged between the PC’s connected to the internet, on a daily basis. The developed technology enables more and more high speed of communication. The data are frequently recorded as time series. We deal with the speed of communication. It is recorded by a PC the speed of information retrieval via the Internet at a frequency of 1 sec. 10.800 values of speed are collected, corresponding to 3 hours. As it is often undertaken that the communication speed is stable, equal values of speed would be expected to be recorded for all the recorded periods of 1 sec. Ultimate measurements indicate that there are huge variations among the values of 10.800 records. The idea of Moving Average (M-A) is used for the 10.800 time series data, with several values of its step parameter k (or depth). As we pass from a time series to its fruit, via the (M-A) process, the dispersion is reduced and the new time series is smoothed, highlighting some trends in behaviors of the initial series.

Key words: Web, Internet, Moving average, Time Series, Communication.

AMS Subject Classification: Primary 62M10, 62M20, 62D05
Secondary 60G25, 93E10, 94A20

1. Introduction
A time series of length \(N\) is a function of the set \(\{1,2,3,4,\ldots,N\}\) to the set of the real numbers and we symbolize it as \(X_1, X_2, X_3, \ldots, X_N\). It is undertaken that \(N\) is a rather big number, e.g. \(N>1000\) or \(N>5000\). Also every element of this series could be seen as a Random Variable (rv) \(X_m\), \(m=1,2,3,\ldots,N\), with the same Variance, \(Var X_m=\sigma^2\). We can undertake also that every rv \(X_m\), \(m=1,2,3,\ldots,N\) is an independent and identically distributed random variable (i.i.d. rv). This is not the law but it is valid for the majority of the cases. We symbolize the above time series as \((x_n)_{n=1}^{N}\), \(n=\)set of indices.

2. Data Description
Our data is a time series with length \(N=10800\) values. They are the values of speed of data transfer from a computer to an other (or from the internet to our personal computer) corresponding to 3 hours of registration, i.e. we have registered the mean value of the speed for a period of 1 second (1 sec), for 10800 sec (=3 hours) of data transfer. The computer was connected to the internet via an ADSL connection with max speed of 24Mbits/sec. The min the max and the average speed values were equal to 850, 1120 and 985.351 Kbits/sec, respectively, with standard deviation \(s=78.08\) and \(Variance=6096.1228\). See the next diagram 2.1.
From a first point of view the data are dispersed in a wide field even if someone waited to select a constant speed for all the 10800 observations. Anyway the data look like randomly distributed uniform data. In a next paragraph we are going to see the behavior of the data and especially their variance via the idea of the mechanism of the moving average (M-A) applied on the above data.

3. Moving Average for labeled data
Suppose we have a series of \( N \) labelled data values \( A_1, A_2, A_3, \ldots, A_N \). (3.1)

We see the next series of values coming out from (3.1):

\[
B_i = \frac{1}{k} \sum_{j=1}^{k} A_{i+j-1}, \quad i = 1, 2, 3, \ldots, N - k + 1
\]  
(3.2)

**Definition 3.1:** The series (3.2) is the Moving Average of (3.1) with step \( k > 1 \). We will call the series (3.1) **Generator**.

Generally we can get a \((M-A)\) from (3.2) with step \( m \) equal or no to \( k \) and so on. The basic offer of the idea of \((M-A)\) to the data analysis is based on the variance of the series data. It will be shown that the variance of the elements of \((M-A)\) is less than the suitable of its generator.

We can see the elements of the series (3.1) as rv independently with mean value \( EA_i = \mu_i \), variance \( Var(A_i) = \sigma_i^2 \), \( i = 1, 2, 3, \ldots, N \). Also the covariance of any two rv \( A_i \) and \( A_m \), the \( Cov(A_i, A_m) = E(A_i \cdot A_m) - E(A_i) \cdot E(A_m) = \sigma_{j,m} \) is a well known idea.

For the series in (3.1) we adopt the symbolism

\[
\sigma_i^2 = E\sigma_i^2 \tag{3.3}
\]

and for \( i = 1, 2, 3, \ldots, N-k+1 \) the one

\[
\sigma_{j,m} = ECov(A_j, A_m), \quad j = i, i+1, \ldots, i+k-2, \quad j < m < i + k - 1 \tag{3.4}
\]
Lemma 3.1: For the (M-A) in (3.2) it is
\[ EB^2_i = \frac{\sum_{u=i}^{i+k-2} E(A)^2_u + 2 \cdot \sum_{j=m}^{i+k-2} E(A_j \cdot A_m)}{k^2}\]
and \( (E_b)^2 = \frac{\sum_{u=i}^{i+k-2} \mu_u^2 + 2 \cdot \sum_{j=m}^{i+k-2} \mu_j \cdot \mu_u}{k^2}. \)

Proof: From (3.2) we have \( EB_i = \frac{1}{k} \sum_{j=m}^{i+k-1} E(A)_{i+j} = \frac{1}{k} \sum_{j=m}^{i+k-1} \mu_{i+j}, \ i = 1, 2, 3, \ldots, N - k + 1 \)
and so it is
\[ (EB)^2_i = \frac{1}{k^2} \left[ \sum_{j=1}^{k} (E(A)_{i+j})^2 + 2 \cdot \sum_{j=1}^{k} \sum_{m=j}^{i+k-1} (E(A)_{i+j}) \cdot (E(A)_{i+m}) \right], \ i = 1, 2, \ldots, N - k + 1 \]
or
\[ (EB)^2_i = \frac{1}{k^2} \left[ \sum_{j=1}^{k} (\mu_{i+j})^2 + 2 \cdot \sum_{j=1}^{k} \sum_{m=j}^{i+k-1} (\mu_{i+j}) \cdot (\mu_{i+m}) \right], \ i = 1, 2, \ldots, N - k + 1. \]

Also from (3.2)
\[ B^2_i = \frac{1}{k^2} \left[ \sum_{j=m}^{i+k-1} A^2_{i+j} + 2 \cdot \sum_{j=m}^{i+k-1} \sum_{m=j}^{i+k-1} A_j \cdot A_m \right], \ i = 1, 2, \ldots, N - k + 1 \]
is obtained and so
\[ EB^2_i = \frac{1}{k^2} \left[ \sum_{j=m}^{i+k-1} A^2_{i+j} + 2 \cdot \sum_{j=m}^{i+k-1} \sum_{m=j}^{i+k-1} (E(A_j \cdot A_m)) \right], \ i = 1, 2, \ldots, N - k + 1, \text{ q.d.e.} \]

Theorem 3.1: For the (M-A) in (3.2) and for \( i=1, 2, 3, \ldots, N-k+1 \) it is
\[ Var_B_i = \frac{\sigma^2 + (k - 1) \cdot \sigma_{jm}}{k}, \ j = i, \ldots, i + k - 2, \ j < m < i + k - 1. \]

Proof: From the idea of variance and for (3.2) we have
\[ Var_B_i = E(B^2_i) - (EB^2_i) \quad (3.5) \]
and from lemma 3.1 we get
\[ Var_B_i = \frac{1}{k^2} \left[ \sum_{u=i}^{i+k-2} \sigma_u^2 + 2 \cdot \sum_{j=m}^{i+k-2} \sum_{m=j}^{i+k-2} (E(A_j \cdot A_m)) \cdot (E(A_u)) \right], \]
or
\[ Var_B_i = \frac{1}{k^2} \left[ \sum_{u=i}^{i+k-2} \sigma_u^2 + 2 \cdot \sum_{j=m}^{i+k-2} \sum_{m=j}^{i+k-2} \sigma_{u,m} \right] = \frac{1}{k^2} \left[ k \cdot \sigma^2 + (k - 1) \cdot \sigma_{j,m} \right] \]
and finally
\[ Var_B_i = \frac{1}{k^2} \left[ \sigma^2 + (k - 1) \cdot \sigma_{j,m} \right] \quad \text{q.d.e.} \quad (3.6) \]
Theorem 3.2: For any two rv \( X \) and \( Y \) and their variances and covariance it is valid:

“"The mean value of the two variances is bigger than the covariance of \( X \) and \( Y \).""

Proof: From the ideas of variance and covariance we get

\[
\text{Var}(X) = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - n \cdot \mu_x^2 = \sigma_x^2 \quad \text{and} \quad \text{Cov}(X,Y) = \frac{1}{n} \sum_{i=1}^{n} x_i y_i - n \cdot \mu_x \cdot \mu_y = \sigma_{xy}
\]

and so for the mean value of the two variances it is

\[
\bar{\sigma}^2 = \frac{\sigma_x^2 + \sigma_y^2}{2} = \frac{1}{2} \sum_{i=1}^{n} x_i^2 + y_i^2 - \left( \mu_x^2 + \mu_y^2 \right)
\]

and after this

\[
\bar{\sigma}^2 - \sigma_{xy} = \frac{\sigma_x^2 + \sigma_y^2 - 2 \cdot \sigma_{xy}}{2} = \frac{1}{2} \left\{ \sum_{i=1}^{n} x_i^2 + y_i^2 - \left( \mu_x^2 + \mu_y^2 \right) \cdot \sum_{i=1}^{n} x_i y_i + 2 \cdot \mu_x \cdot \mu_y \right\}
\]

or

\[
\bar{\sigma}^2 - \sigma_{xy} = \frac{1}{2} \text{Var}(X - Y) \geq 0
\]

and finally \( \bar{\sigma}^2 \geq \sigma_{xy} \), q.d.e.

Corollary 3.1: “The mean value of the two variances is bigger than the mean covariance of \( X \) and \( Y \).”

Proof: Since there is only one value for the covariance this value is the mean value too. So from theorem 3.2 we get \( \bar{\sigma}^2 \geq \sigma_{xy} \).

Theorem 3.3: Suppose we face a group of \( k \) rv, the \( X_1, X_2, X_3,..., X_k \) and the suitable \( k \) variances \( \sigma_1^2, \sigma_2^2,..., \sigma_k^2 \) and \( k(k-1)/2 \) covariances \( \sigma_{j,i}, i=1,2,3,...,k-1, \ i < j < k \). We adopt the symbol \( \bar{\sigma}^2 = E\sigma_i^2 \) for the mean value of the \( k \) variances and the symbol \( \bar{\sigma}_{j,m} = E\text{Cov}(X_j,X_m), j=1,2,...,k-1, \ j < m < k \), for the mean value of the \( k(k-1)/2 \) covariances. The next statement is valid:

“The mean value of the \( k \) variances is bigger than the mean value of the \( k(k-1)/2 \) covariances, \( \bar{\sigma}_{j,m} = E\text{Cov}(X_j,X_m), j=1,2,...,k-1, \ j < m < k \), i.e. \( \bar{\sigma}^2 \geq \bar{\sigma}_{j,m} \).

Proof: From theorem 3.2 and for any pair of rv \( X_j, X_m \), \( j=1,2,...,k-1, \ j < m < k \) we have the validity of

\[
\sigma_j^2 + \sigma_m^2 > 2 \cdot \sigma_{j,m} , j=1,2,...,k-1, \ j < m < k \quad (3.7)
\]

In the system of (3.7) there are \( k(k-1)/2 \) equations and the variance \( \sigma_i^2, i=1,2,...,k \) appears \( k-1 \) times. Let us see on the summation
\[
\sum_{j=1}^{k-1} \sum_{j=1}^{k} \left( \sigma^2_j + \sigma^2_m \right) > 2 \sum_{j=1}^{k-1} \sum_{j=1}^{k} \sigma_{j,m} \quad \text{or}
\]
\[
(k - 1) \cdot \sum_{i=1}^{k} \sigma^2_i = (k - 1) \cdot k \cdot \sigma^2 > 2 \sum_{j=1}^{k-1} \sum_{j=1}^{k} \sigma_{j,m}
\]
or finally dividing the three parts of the above relation by \(k(k-1)\) we obtain:
\[
\frac{2 \sum_{j=1}^{k-1} \sum_{j=1}^{k} \sigma_{j,m}}{k(k-1)} = \bar{\sigma}_{j,m}, \quad \text{q.d.e.}
\]

4. Comparing the variances and covariances

The \((M-A)\) has a good and utilizable property expressed by the next theorem: **Theorem 4.1:** Suppose we have the series in (3.1) and its \((M-A)\) in (3.2). Step = \(k\).

The variance of the rv \(B\) in (3.2) is less than the mean variance of \(A\), in (3.3).

**Proof:** From theorems 3.1 and 3.3 we conclude that
\[
\text{Var}B_j = E B_j^2 - \left( E B_j \right)^2 = \frac{1}{k} \left( \sigma^2_1 + (k - 1) \cdot \bar{\sigma}_{j,m} \right) < \frac{1}{k} \left( \sigma^2_1 + (k - 1) \cdot \sigma^2 \right) = \bar{\sigma}^2 \quad \text{q.d.e.}
\]

**Corollary 4.1:** If in (3.1) \(A_i, i=1, \ldots, N\) are i.i.d., then \(\text{Var}B_j = E B_j^2 - \left( E B_j \right)^2 = \frac{\sigma^2}{k}\).

**Proof:** Assuming that \(A_i\) are i.i.d. we get that \(\bar{\sigma}_{j,m} = 0\), \(j = 1, 2, \ldots, k - 1\), and
\[
j < m < k.
\]

**Corollary 4.2:** If in (3.1) \(A_i, i=1, 2, \ldots, N\) are i.i.d., then the standard deviation of the \((M-A)B_j, j=1, 2, \ldots, N-k+1\) is given \(s = \sqrt{\text{Var}B_j} = \frac{\bar{\sigma}}{\sqrt{k}}\).

**Proof:** Obvious from corollary 4.1.

We adopt for the \((M-A)B_j, j=1, 2, \ldots, N-k+1\) the notation \(\text{Var}B_j = \bar{\sigma}^2 \) and continuously we have to write \(\text{Var}A_i = \bar{\sigma}^2 \) . So a simple algorithm testing if the \(A_i, i=1, 2, \ldots, N\) are really i.i.d., could be the next based on theorem 3.1 and on corollaries 4.1, 4.2:

1\textsuperscript{st}) Solve equation in theorem 3.1 for the covariance and get the solution
\[
\bar{\sigma}_{j,m} = \frac{k \cdot \bar{\sigma}^2 \cdot (k - 1)}{k - 1} \quad \text{(4.1)}
\]

2\textsuperscript{nd}) Apply on (4.1) and find the value of \(\bar{\sigma}_{j,m} \). Check

If \(\bar{\sigma}_{j,m} = 0\), then \(A_i, i=1, 2, \ldots, N\) are i.i.d.
Else if $\sigma_{j,m} \neq 0$, then $A_i, i=1, 2, ..., N$ are not i.i.d. and, among them, there will be some kind of correlation. The majority of the cases give $\sigma_{j,m} \equiv 0$ instead of $\sigma_{j,m} = 0$. Also an alternating sign of $\sigma_{j,m}$ for contiguous values of $k$ means that $\sigma_{j,m} = 0$ and if the absolute value of $\sigma_{j,m}$ is a 10% or 15% of $\sigma^2(k)$, then the $\sigma_{j,m} = 0$ is also recognized.

See for examples below. The result of the theorem 4.1 enables anyone to substitute the series (rv) $A$ in (3.1) with its $(M-A)$, the rv $B$ in (3.2). The rv $B$ has less variance than the initial one $A$ and offers a smoothing to the values of $A$. This effect gives also a good figure for the trends of the values of rv $A$ via the values of the rv $B$. It is valid that the bigger the step $k$ the stronger the influence of smoothing as we can see also in example 5.2. This smoothing makes temporary trends to be vanishing. This will be clearer by the examples based on the data mentioned in paragraph 2.

5. Examples

Some illustrative examples will be presented in this paragraph:

Example 5.1: The data described in paragraph 2 gave for step values $k=1,2,\ldots,100$ results for $\sigma^2(k)$. We applied the formula (4.1) for the corresponding values of $\sigma_{j,m}$. The first 7 cases and the last 7 of these, figure in the Table 5.1:

<table>
<thead>
<tr>
<th>k</th>
<th>$\sigma^2(k)$</th>
<th>$\sigma_{j,m}$</th>
<th>i.i.d. or no</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6096,12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3031,70</td>
<td>-32,72</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>2030,07</td>
<td>-2,95</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>1526,25</td>
<td>2,96</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>1220,54</td>
<td>1,64</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>1020,67</td>
<td>5,57</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>874,40</td>
<td>4,11</td>
<td>Yes</td>
</tr>
<tr>
<td>94</td>
<td>58,49</td>
<td>6,43</td>
<td>Yes</td>
</tr>
<tr>
<td>95</td>
<td>57,82</td>
<td>6,41</td>
<td>Yes</td>
</tr>
<tr>
<td>96</td>
<td>57,19</td>
<td>6,38</td>
<td>Yes</td>
</tr>
<tr>
<td>97</td>
<td>56,58</td>
<td>6,34</td>
<td>Yes</td>
</tr>
<tr>
<td>98</td>
<td>55,97</td>
<td>6,30</td>
<td>Yes</td>
</tr>
<tr>
<td>99</td>
<td>55,39</td>
<td>6,25</td>
<td>Yes</td>
</tr>
<tr>
<td>100</td>
<td>54,81</td>
<td>6,21</td>
<td>Yes</td>
</tr>
</tbody>
</table>

There is very little values for the covariance and in the first part the sign is changeable. Only the first and the last 7 cases are presented. For the other cases similar scores are detected.
Example 5.2: The data described in paragraph 2 and for step values \(k = 1, 2, \ldots, 100\) gave also results for \(\sigma(k)\) and \(\ln \sigma(k)\). We have put them in the next Table 5.2.

For space reasons again in this table only 14 cases are presented, the first 7 and the last 7, even if the results are for all the 100 cases:

<table>
<thead>
<tr>
<th>(k)</th>
<th>(\sigma(k))</th>
<th>(\ln \sigma(k))</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>78.08</td>
<td>4.357704</td>
<td>(\ln \sigma(k) = 3.26 - 0.0148 \cdot k)</td>
</tr>
<tr>
<td>2</td>
<td>55.06</td>
<td>4.008440</td>
<td>With correlation coefficient (r = -0.902155)</td>
</tr>
<tr>
<td>3</td>
<td>45.06</td>
<td>3.807914</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>39.07</td>
<td>3.665286</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>34.94</td>
<td>3.553524</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>31.95</td>
<td>3.464105</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>29.57</td>
<td>3.386770</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>28.22</td>
<td>3.314472</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>26.87</td>
<td>3.244263</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>25.53</td>
<td>3.175185</td>
<td></td>
</tr>
<tr>
<td>……</td>
<td>……</td>
<td>……</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>7.65</td>
<td>2.034448</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>7.60</td>
<td>2.028696</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>7.56</td>
<td>2.023184</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>7.52</td>
<td>2.017791</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>7.48</td>
<td>2.012433</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>7.44</td>
<td>2.007164</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>7.40</td>
<td>2.001955</td>
<td></td>
</tr>
</tbody>
</table>

In the second column we have the values of standard deviation of \(A_i\), \(i = 1, 2, \ldots, N\), (1st line, \(k = 1\)) and for \(B_i(k), k = 2, \ldots, 7\) (lines 2nd, 3rd, …, 7th) and \(k = 94, 95, \ldots, 100\) (last 7 lines). In the third column we have the Neper logarithm of the previous.

1st) We have plotted the data of the 2 fist columns of Table 5.2. We can describe the formula of the plotted data as

\[
\sigma(k) = \frac{\bar{\sigma}}{\sqrt{k}} + e \quad (5.1)
\]

where \(e\) = error due to \(\sigma_{j,m} \approx 0\) instead of \(\sigma_{j,m} = 0\) (we can note \(e \approx \sqrt{\sigma_{j,m}}\)).

Diagram 5.1

Connecting the values of the standard deviation of \((M-A) B(k)\) and the value of the step \(k\) via the mean standard deviation of series \(A\).
2nd) We tried to get a regression line connecting the values of $\ln \sigma(k)$ (3rd column of Table 5.2) and the values of the step $k$ of (M-A) (1st column of Table 5.2). Since the correlation coefficient is $r = -0.90215$, a 80% of variation of $\ln \sigma(k)$ of the (M-A) depends on the value of the step (depth) of the (M-A). So we can state:

“The bigger the step $k$ the stronger the influence of smoothing”.

The regression line is $\ln \sigma(k) = 3.262574 - 0.014819 \cdot k$

![Diagram 5.2](image)

**Diagram 5.2**

Connecting the values of the Neper logarithm standard deviation of (M-A) $B(k)$ and the value of the step $k$ via the mean standard deviation of series $A$.

**References**

ULTIMUM ARGUMENT SUPREMUM FOR ASSET PRICE 
MODELLED BY GAUSSIAN SUBORDINATED PROCESSES

STERGIOS FOTOPOULOS AND YUXING LUO
Department of Finance & Management Science, Washington State University

VENKATA JANDHYALA
Department of Mathematics, Washington State University.

The aim of this article is to develop a theory concerning the ultimum argument supremum of subordinated Gaussian processes with negative drift. This includes obtaining exact Laplace transform of the argument supremum, its density and its stochastic representation. The stochastic representations show that the ultimum argument supremum admits the Wiener-Gamma integrals wherein only the underlying distribution may change whereas the parameter remains constant throughout all subordinators. Considering various examples, the theoretical results for the general case have been further investigated for some of the well-known subordinators. Numerical investigations of distribution of the argument supremum show close agreement between the theoretical derivations and empirical computations. The Barndorff-Nielsen and Shephard model, which has received much attention in the recent financial literature, is the central motivation for considering the Gaussian subordinated process in this article.

KEY WORDS: Wiener-Hopf factorization, supremum of Lévy process, Spitzer’s condition, linear Brownian motion, Gaussian subordinated process.

Address Correspondence to Stergios B. Fotopoulos, Department of Finance & Management Science, Washington State University, Pullman, WA 99164-4746, USA; email: fotopo@wsu.edu.
1. INTRODUCTION

Some of the key results regarding the Brownian motion that found wide ranging applications not only in finance but also in all areas of science and technology are supremum of the Brownian motion, and the arc sine law, which capture the distribution of the argument supremum normalized by the time until the process has been observed. At a technical level, the arc sine law holds whenever the underlying Brownian motion visits the zero horizontal line infinitely often (oscillating) (e.g., see Feller 1971 II, Spitzer 1976, Bertoin 1996, Sato 1999, Karatzas and Shreve 1988). On the other hand, when a negative (positive) drift is involved in the Brownian motion process, one may still be interested in deriving a limiting result for the distribution of the argument supremum (infimum). For example, in the presence of a negative drift, it turns out that the ultimate supremum is finite (see Sato 1999). Consequently, the argument of the supremum is also finite and one can study the limiting distribution of the argument supremum without the need for a normalizing constant. The main aim of this article is to characterize the limiting distribution of the argument supremum of Gaussian subordinated processes when the subordinator has both killing rate and drift equal to zero.

Before we move on to the Gaussian subordinated processes, we first provide some general results related to the argument maximum for Lévy processes. The fundamental approach we employ in these derivations stems from the Wiener-Hopf factorization and related identities on random walks and Lévy processes. Historically, these identities are a result of many researchers that include among others Spitzer (1976), Feller (1971), Pecherskii and Rogozin (1969), Bingham (1975), Gusak and Korolyuk (1974), Greenwood (1975), Greenwood and Pitman (1980), Bertoin (1996), Sato (1999), Doney (2005), and Kyprianou (2004). The importance of the Wiener-Hopf
factorization for either random walks or Lévy processes is that it characterizes the range of the processes running maximum as well as the times at which new maxima are attained. Subsequently, taking advantage of some of the characterization results in Bondesson (1992), James et al. (2008a) and James et al. (2008b), we have derived explicit characterizations for the limiting distribution of the argument supremum of the Brownian motion with negative drift and also of Gaussian subordinated processes.

From an applications perspective, the central role of Brownian motion processes and more generally, Lévy processes of which Gaussian subordinated processes are special cases in modeling financial log-returns is well documented in the literature (see e.g., Barndorff-Nielsen et al. 2001 and the references therein). There are well known limitations to how well Gaussian processes can be applied to model log-returns of high frequency assets (see Eberlein 2001). For example, empirical properties of high frequency log-returns such as heavy tails and jumps could not be accommodated under the usual Gaussian processes framework, but could be suitably accommodated through the Lévy processes. In particular, Gaussian subordinated processes have been found to be suitable for modeling high frequency log-returns (see Eberlein 2001). Results regarding the distribution of argument supremum of such processes are derived for the first time in this article. In particular, our derivations show that the argument supremum defined upon the simple linear Brownian motion process as well as the subordinated Gaussian process can be expressed as appropriate Wiener-Gamma integrals with a constant parameter 0.5. These results can be expected to play a definitive role in understanding the behavior of financial log-returns. For example, the argument supremum can be identified with the time at which log-returns attain maximum in a period of declining drift or the argument of the infimum in a period of sustained growth in the market.
The contents of this investigation are organized as follows: Section 2 introduces the Lévy process and presents some general results on fluctuation theory in continuous time governed by the Lévy process. Specifically, it bridges results from discrete to continuous time by showing various identities of the discrete analogs. Section 3 specializes to the subordinated Gaussian process and presents the main results of this study. Section 4 provides proofs of the main results and several examples of the ultimum argument supremum are presented in Section 5 for well-known subordinators. We display graphs of the proposed theoretical results as well as through empirical investigations in Section 6 and end the article with some concluding remarks in Section 7.

2. SOME ELEMENTS FOR LADDER EPOCHS ON LÉVY PROCESSES

In studying fluctuation theory of Lévy processes, the Wiener-Hopf factorization plays an important role in deriving and analyzing various properties such as excursions and first passage times. In this section, we work with an expression linked to the first (last) hitting time that a process reaches its maximum, when the underlying process is Lévy. Assuming that a process drifts to $-\infty$, we are interested to recover the probability distribution of this index for various indices of Lévy processes. We refer the reader to Sato (1999, Ch. 9) and Bertoin (1996) for a detailed discussion on the subject of a Lévy process.

Let $\{X(t): t \geq 0\}$ be an $\mathbb{R}$-valued Lévy process defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, where the filtration $\{\mathcal{F}_t\}$ satisfies the usual conditions of right continuity and completion. We take $P(X_0 = 0) = 1$. It is well known from the Lévy-Khintchine formula that a Lévy process on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ satisfies $E[\exp(i\theta X(t))] = e^{\psi(\theta)}$, $t \geq 0$, $\theta \in \mathbb{R}$, where
GAUSSIAN SUBORDINATED PROCESS AND ITS ARGUMENT SUPREMUM

(2.1) \[ \psi(\theta) = i\mu \theta - \frac{1}{2} \sigma^2 \theta^2 + \int_{-\infty}^{\infty} \left( 1 - e^{i\theta x} + ix \phi(|x| < 1) \right) \nu(dx), \]

with \( \mu \in \mathbb{R} \), \( \sigma \geq 0 \), and \( \nu \) on \( \mathbb{R} \setminus \{0\} \) is the Lévy measure that satisfies \( \int_{-\infty}^{\infty} x^2 \wedge 1 |\nu(dx)| < \infty \).

The function \( \psi \) defined for \( \Re(\theta) \geq 0 \) is referred to as a Lévy exponent. Its restriction to the non-negative real line is strictly convex and satisfies \( \lim_{\theta \to \infty} \psi(\theta) = \infty \). In this article, we shall denote a Lévy process by the triplet \((\mu,\sigma,\nu)\).

Let \( \overline{X}(t) := \sup_{s \leq t} X(u) \) and \( \underline{X}(t) := \inf_{s \leq t} X(u) \) denote the supremum and the infimum of the processes \( X_t \), \( t \geq 0 \), respectively. It is well known, that the excursions away from the maximum can also be viewed as those away from zero of the reflected process \( R(t) := \overline{X}(t) - X(t) \). Let \( \tau \) be an exponential random variable with parameter \( q > 0 \), which is independent of the process \( \{X(t) : t \geq 0\} \). Further, define the first passage times to maximum or minimum by

\[ T^+_x := \inf \{ t > 0 : X(t) > x \} \quad \text{and} \quad T^-_x := \inf \{ t > 0 : X(t) \leq x \}, \quad \text{for} \ x \in \mathbb{R}. \]

When \( x = 0 \), \( T^+_0 \) and \( T^-_0 \) present the first ascending and first descending ladder epochs, respectively. On the same lines, we also introduce the pairs of indices:

\[ J^+_i := \inf \{ s < t : \overline{X}(t) = X(s) \} \quad \text{and} \quad J^-_i := \inf \{ s < t : \underline{X}(t) = X(s) \}, \]
\[ G^+_i := \sup \{ s < t : \overline{X}(s) = X(s) \} \quad \text{and} \quad G^-_i := \sup \{ s < t : \underline{X}(s) = X(s) \}. \]

The aim of the article is to understand the behavior of \( J^+_i \), the index at which the process reaches its maximum for the first time by time \( t \). Note that upon proper standardization, \( J^+_i \) converges to the arc-sine law. Here, upon allowing the Lévy process to drift to \(-\infty\), we aim to determine the distribution of \( J^+_i \) as \( t \) tends \( \infty \). It can be shown that when the process drifts to
\(-\infty\) (negative spectrally Lévy process), \(J^+_\infty\) is almost surely bounded. Hence, it is important to derive its distribution without any further standardization. Restricting to the process \(J_t^+, t \geq 0\), we first obtain an identity similar to that of Pecherskii and Rogozin (1969).

Noting that \(\{X(t) : t \geq 0\}\) is right continuous with finite left limits, we let \(n \in \mathbb{N} \setminus \{0\}\), and discretize the process by

\[
X^n_k(t) := X_k/2^n, \quad \text{for} \quad t \in \left[ \frac{k}{2^n} \cdot \frac{k + 1}{2^n} \right], \quad k = 0, 1, \ldots, 2^n - 1
\]

with \(X^n_k(0) = X_0 = 0\). Evidently, the constructed map \((k, \omega) \mapsto X^n_k(\omega)\) from \([0, \infty) \times \Omega\) into \(\mathbb{R}\) is \(\mathcal{B}\left([0, \infty)\right) \otimes \mathcal{T}_t\)-measurable. By right-continuity, one obtains that \(\lim_{n \to \infty} X^n_k(t) = X(t)\) for any \(t \in \left[ \frac{k}{2^n} \cdot \frac{k + 1}{2^n} \right]\). Thus, the map from \(t \mapsto X(t)\) is clearly \(\mathcal{B}\left([0, \infty)\right) \otimes \mathcal{T}_t\)-measurable.

It is known from the discrete case (see, e.g., Feller VII, 1971, p.416) that

\[
P(s) = \sum_{n=0}^{\infty} s^n p_n = \exp \left( \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n > 0) \right) \quad \text{and}
\]

\[
Q(s) = \sum_{n=0}^{\infty} s^n q_n = \exp \left( \sum_{n=1}^{\infty} \frac{s^n}{n} P(S_n \leq 0) \right), \quad s \in (0,1),
\]

where \(P(s)\) and \(Q(s)\) represent probability generating functions of the corresponding sequences \(p_n = P\left(\cap_{j=1}^{n} (S_j > 0)\right), \quad q_n = P\left(\cap_{j=1}^{n} (S_j \leq 0)\right)\), respectively and \(S_j, \quad j \geq 1\), denotes partial sums of i.i.d. variables. Let \(J^+_n, \quad n \geq 1\), denote the discrete analogue of \(J_t^+, \quad t \geq 0\). It can be seen that

\[
P(J^+_n = j) = p_j q_{n-j}, \quad j = 0, 1, \ldots, n, \quad n \geq 1,
\]
where \( P_n \) and \( q_n \) may equally be seen as \( p_n = P(T_0^- > n) \) and \( q_n = P(T_0^+ > n) \), \( n \geq 1 \). It is then clear that for \( s, x \in (0,1) \),

\[
(1-x)\sum_{n=0}^{\infty} x^n E[s^{J_n}] = (1-x)\sum_{n=0}^{\infty} x^n \sum_{j=0}^{n} s^j p_j q_{n-j}
\]

(2.4)

\[
= (1-x)\sum_{j=0}^{\infty} (sx)^j p_j \sum_{n=j}^{\infty} x^{n-j} q_{n-j} = (1-x)P(sx)Q(x)
\]

\[
= \exp \left( -\sum_{n=1}^{\infty} \frac{x^n}{n} (1-s^n)P(S_n > 0) \right).
\]

In light of the basic Lévy properties for stochastic processes, we set \( x = e^{-q/2^n}, s = e^{-\nu/2^n} \), and \( S_k \equiv X_k^n(t) \) for \( t \in \left[ \frac{k}{2^n}, \frac{k+1}{2^n} \right] \) and \( n \in \mathbb{N} \setminus \{0\} \), and \( k = 0, 1, \ldots, 2^n - 1 \). Upon letting \( n \uparrow \infty \), the left-hand-side of (2.4) tends to the following limit:

\[
(1-x)\sum_{n=0}^{\infty} x^n E[s^{J_n}] \rightarrow q \int_0^{\infty} e^{-qu} du E[e^{-\nu J_k^n/2^n}].
\]

(2.5)

Similarly, letting \( n \uparrow \infty \), the right-hand-side of (2.4) converges to:

\[
\exp \left( -\sum_{n=1}^{\infty} \frac{x^n}{n} (1-s^n)P(S_n > 0) \right) = \exp \left( -\sum_{n=1}^{\infty} \frac{e^{-q/2^n}}{\left( k/2^n \right)} \frac{1}{2^n} E[e^{-\nu J_k^n/2^n}] \right)
\]

(2.6)

Combining (2.5) and (2.6), the above results yield the following theorem:
THEOREM 2.1. Let \( \{X_t : t \geq 0\} \) be a real-valued Lévy process and let \( \tau \) be exponentially distributed with parameter \( q > 0 \), independent of the process. Then,

\[
E[\exp(-\lambda J_\tau^+)] = \exp\left\{ -\int_0^\infty \frac{e^{-qu}}{u} (1 - e^{-\lambda u})P(X(u) > 0)\right\}, \quad \lambda > 0.
\]

The Laplace transform of \( G_\tau^+ \) derived by Greenwood and Pitman (1980) is identical to the expression in the above theorem, and hence both \( G_\tau^+ \) and \( J_\tau^+ \) have the same law.

Upon noting that \( q\int_0^\infty e^{-qu} du E[e^{-\lambda J_\tau^+}] = \int_0^\infty e^{-u} du E[e^{-\lambda J_\tau^+}] \to E[e^{-\lambda J_\tau^+}] \) as \( q \to 0 \), the following theorem follows:

THEOREM 2.2. Let \( \{X_t : t \geq 0\} \) be a real-valued Lévy process. Then, for \( \lambda > 0 \), we have:

\[
E[\exp(-\lambda J_\tau^+)] = \exp\left\{ -\int_0^\infty \frac{du}{u} (1 - e^{-\lambda u})P(X(u) > 0)\right\}.
\]

In a seminal paper, Greenwood and Pitman (1980) showed that for a random variable \( \tau \sim \text{Exp}(q) \), which is independent of the process \( \{X(t): t \geq 0\} \), the bivariate process \( (\tau, X(\tau)) \) can be decomposed as \( (\tau, X(\tau)) = (G_\tau^+, \overline{X(\tau)}) + (\tau - G_\tau^+, X(\tau) - \overline{X(\tau)}) \), i.e., the supremum and its reflection. Note that \( \tau \sim \text{Exp}(q) \) is infinitely divisible with no drift and with Lévy measure given by \( t^{-1}e^{-qt} dt \), they have also shown that the random vector \( (\tau, X(\tau)) \) has infinitely divisible distribution with Lévy measure given by \( \mu(dt, dx) = t^{-1}e^{-qt} P(X(t) \in dx) dt \). Using time reversal, it follows that

\[
(\tau - G_\tau^+, X(\tau) - \overline{X(\tau)}) =_d (G_\tau^-, \overline{X(\tau)}).
\]

To complete the theme initiated above, one can argue as in equation (2.7) to obtain similar results as in Theorems 1 and 2 for the reversal process. Specifically, we have:
Corollary 2.3. Let \( \{X(t) : t \geq 0\} \) be a real-valued Lévy process and let \( \tau \) be exponentially distributed with parameter \( q > 0 \), independent of the process. Then,

\[
E[\exp(-\lambda J^-_\tau)] = \exp\left( -\int_0^\infty \frac{e^{-\nu u}}{u} \left( 1 - e^{-\lambda u} \right) P(X(u) < 0) \right), \quad \lambda > 0.
\]

Upon letting \( q \to 0 \), one may also obtain the next result.

Corollary 2.4. Let \( \{X(t) : t \geq 0\} \) be a real-valued Lévy process. Then, for \( \lambda > 0 \), we have:

\[
E[\exp(-\nu J^-_\tau)] = \exp\left( -\int_0^\infty \frac{du}{u} \left( 1 - e^{-\nu u} \right) P(X(u) < 0) \right).
\]

Many important consequences follow from Theorems 1 and 2, which we state without proof.

Theorem 2.5. Let \( \{X(t) : t \geq 0\} \) be a real-valued Lévy process. Then,

i. If \( \int_1^\infty \frac{du}{u} P(X(u) > 0) < \infty \), then \( \lim_{t \to \infty} X(t) < \infty \) a.s. and \( \lim_{t \to \infty} X(t) = -\infty \) a.s.

ii. If \( \int_1^\infty \frac{du}{u} P(X(u) < 0) < \infty \), then a.s. \( \lim_{t \to \infty} X(t) = \infty \) and \( \lim_{t \to \infty} X(t) > -\infty \) a.s.

iii. If \( \int_1^\infty \frac{du}{u} P(X(u) > 0) < \infty \), then \( P(J^+_\infty < \infty) = 1 \) and \( P(J^-_\infty = \infty) = 1 \).

iv. If \( \int_1^\infty \frac{du}{u} P(X(u) < 0) < \infty \), then \( P(J^+_\infty = \infty) = 1 \) and \( P(J^-_\infty < \infty) = 1 \).

In the above, Bertoin (1996) has earlier seen (i) and (ii) while (iii) and (iv) are perhaps new.

Before we continue with the random variable \( J^+_\infty \), it is worth noting the differences between the asymptotic law of \( J^+_\infty / t \) and the law of \( J^+_\infty \). Contingent on Spitzer’s condition (Bertoin, 1996):

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t P(X(u) > 0) du = \rho \in [0, 1],
\]

it has been shown that \( J^+_\infty \) follows the generalized arc-sine law concentrated on \([0,1]\) given by
GAUSSIAN SUBORDINATED PROCESS AND ITS ARGUMENT SUPREMUM

\[(2.9) \quad \lim_{t \to \infty} P \left( \frac{J_t^+}{t} \in du \right) = \frac{(u)^{\rho-1}(1-u)^{-\rho}}{\Gamma(\rho)\Gamma(1-\rho)} du = \frac{\sin \rho \pi}{\pi} (u)^{\rho-1}(1-u)^{-\rho} du, \quad u < 1. \]

Bertoin (1996) has also demonstrated that the results remain the same if the limits in either (2.8) or (2.9) are replaced by \( t \downarrow 0 \). It should be noted that equation (2.9) holds if one replaces \( t^{-1}J_t^+ \) either by \( t^{-1}G_t^+ \), \( t > 0 \) or \( \frac{1}{t} \int_0^t I(X(u) > 0) du \). Moreover, when \( \rho = 0 \) (or 1), the Dirac point mass at zero (mass at 1) replaces the density.

From Theorem 2.5, we note the distribution of \( J_\infty^+ \) depends upon \( P(X(u) > 0), \quad u > 0 \). To capture the behavior of \( J_\infty^+ \) further, one may explore \( J_\infty^+ \) through various forms of the measure \( \mu_t(\,dt,0,\infty) = t^{-1}P(X(u) > 0) \,dt \). We call that such measures can be expressed as Laplace transforms of wide class of subordinators and they will play a key role in understanding the behavior of the random variable \( J_\infty^+ \).

3. THE ULTIMUM ARGUMENT SUPREMUM UNDER THE BARNDORFF-NIELSEN AND SHEPHARD MODEL

Before we get into analytical developments, we first give some background regarding the role that mixtures of Gaussian processes play in the area of financial economic modeling. In the last two decades, it has been widely shown in the literature that the standard option pricing model of Black and Scholes (1973) and Merton (1973) is inconsistent with option data. Looking at empirical densities of log-returns from financial data, one observes the following stylized features: there is more mass near the origin versus the normal distribution, less at the sides than the normal and
considerably more mass in the tails. This means that tiny price movements occur with higher frequency, small and middle-sized movements with lower frequency, and big changes (high jumps) are much more frequent than predicted by normal distribution. Another point for consideration which seems to fit well with empirical data is the asymmetric behavior of the log-returns. Incorporating the above observations, Carr et al. (2003) and Geman et al. (2001) suggest that price processes for financial assets must have a jump component but may not have a diffusion component. Their argument rests on the recognition that all price processes of interest may be regarded as Brownian motion mixed to a random clock. These types of models do support the three points mentioned earlier. To this end, modifying the linear Brownian motion model of Black and Scholes (1973) and Merton (1973), recent literature recommends that one may work with the model:

\[(3.1)\]

\[P(t) = P(0)\exp(X(t)), \quad t \geq 0,\]

where \(P_0 > 0\) is the initial value of an asset (usually non-random), and \(\{X(t) : t \geq 0\}\) is a scale mixture of Gaussian process.

To better describe the model, we let \(\{B(t) : t \geq 0\}\) be a Brownian motion and, \(\mu_t\) and \(\sigma_t\) be adapted processes to the filtration \(\{\mathcal{F}_t : t \geq 0\}\). Clearly, the Itô process

\[(3.2)\]

\[X(t) = \int_0^t \sigma_s dB(s) + \int_0^t \left(\mu_s - \frac{1}{2} \sigma_s^2\right) ds,\]

leads to the following stochastic deferential equation, \(dX(s) = \left(\mu_s - \frac{1}{2} \sigma_s^2\right) dt + \sigma_s dB(s)\). In model (3.1), we let \(P(0)\) be a non-random positive quantity. Model (3.2) suggests that the asset price \(P(t), \quad t \geq 0,\) is allowed to have both instantaneous mean rate \(\mu_t\) and volatility \(\sigma_t\). In this matter, both the mean rate and volatility are permitted to be time varying and random. Clearly, when
\( \mu \) and \( \sigma \) are constant, the geometric Brownian motion model, \( P(t), \ t \geq 0 \), has a log-normal distribution. This, of course, is the well-known Black-Scholes-Samuelson model, which is described by the well-known stochastic differential equation: \( dX(t) = (\mu + \delta \sigma^2)dt + \sigma dB(t) \). A more appealing model in financial economics is the Barndorff-Nielsen and Shephard (2001) model, that proposes a continuous time stochastic volatility (SV) described by the following equation
\[
 dX(t) = (\mu + \delta \sigma_i^2)dt + \sigma_i dB(t), \ \mu \leq 0, \ \delta < 0, \ t \geq 0,
\]
where \( X(t) \) (again) denotes the log-price level, \( \sigma_i \) represents instantaneous volatility independent of the Brownian process \( B(t) \) and the other two constants will be revealed at some later point. The induced likelihood model is based on the integrated volatility \( T(t) = \int_0^t \sigma_i^2 ds \) and can be described as follows. Let \( X(t) \) denote the aggregate return of the log-stock price over the interval \([0,t]\). Thus, on \([0,t]\) the stochastic differential equation of \( X(t) \) has a solution expressed in terms of the Itô stochastic integral of \( \sigma_i \) by
\[
 X(t) = \mu t + \delta T(t) + \int_0^t \langle \sigma_i, dB(s) \rangle, \ \mu \leq 0, \ \delta < 0, \ t \geq 0,
\]
where the real-valued and measurable adapted process \( \sigma_i, t \geq 0 \), is a non-negative \( \mathcal{F}_t \)-predictable stochastic process. The process \( T(t) = \int_0^t \sigma_i^2 ds, \ t \geq 0 \), is assumed to possess homogeneity in both time and space and it is defined such that \( P\left( \int_0^t \sigma_i^2 ds < \infty \ \forall \ t \geq 0 \right) = 1 \). Note that \( \int_0^t \langle \sigma_i, dB(s) \rangle \) and \( T^{1/2}(t)N \) have the same law, where \( N \) is a standard normal variable independent of the process \( \{T(t): t \geq 0\} \). Clearly, in such a setting, the time changed process \( Y(t) = B(T(t)) \) has a
mixed normal distribution, i.e., \( Y(t) \mid T(t) \sim N(0, \sigma^2(t)) \), and is a continuous local martingale. Here, \( \sigma_t \) is clearly a Non-Gaussian process. Below, we exploit expressions (suitable for financial economics) of \( T(t) > 0 \), in order to identify corresponding stochastic counterparts of the random variable \( J^+_\infty \).

The main contribution of the article is to examine how the measure

\[
(3.5) \quad \Lambda(du) = E \left[ \frac{du}{u} \Phi \left( \sqrt{2} \sigma T^{1/2}(u) \right) \right]
\]

where \( q = |\delta| / \sqrt{2} \sigma \), plays an important role in understanding the random variable \( J^+_\infty \) when the process \( T(t) \) drifts to \(-\infty\). Note that \( 3.5 \) depends on the process \( \{T(t) : t \geq 0\} \). Motivated by this and the fact that the process \( T \) is additive and nonnegative, we next introduce a comprehensive class of subordinators whose Laplace exponent is \( \psi \).

In order to deduce equation (3.5), we find it necessary to set \( \mu = 0 \). Consequently, upon substituting \( \mu = 0 \) in model (3.4), the general measure \( \Lambda(dr) = \frac{dr}{t} P(X(t) > 0) \) can be seen to yield (3.5). Thus, in what follows, the model under consideration becomes

\[
(3.6) \quad X(t) \overset{(law)}{=} \delta T(t) + T^{1/2}(t) N, \quad \delta < 0, \quad t \geq 0,
\]

where \( N \) is a standard normally distributed random variable.

In light of (3.6), and the fact that the process \( T = \{T(t) : t \geq 0\} \) is a subordinator, the resulting process described by (3.6) is just a Gaussian subordinated process. To this end, the need to investigate properties of subordinators is desirable.

Let \( \{T(t) : t \geq 0\} \) be a subordinator, that is, an increasing Lévy process taking values in \([0, \infty]\) with \( T_0 = 0 \), where infinity serves as a cemetery point (i.e., an absorbing point). Note that \( T \) is a
right-continuous increasing process adapted to the filtration $\mathcal{F}_t$ started from zero. Let the lifetime of $T$ be defined on the interval $[0, \varsigma)$, where $\varsigma = \inf \{ t \geq 0 : T(t) = \infty \}$. If $\varsigma = \infty$ a.s., then $T$ is a subordinator in a strict sense. The law of a subordinator is specified by the Laplace transform of its one-dimensional distribution. Using the convention, $e^{-\varsigma x} = 0$, for any $\vartheta \geq 0$, we then express the Laplace transform by $E[e^{-\vartheta T(u)} u < \varsigma] = E[e^{-\vartheta T(u)}], u \geq 0, \vartheta > 0$. The homogeneity (additivity) of the increments then yields that the multiplicative property is satisfied. This, in turn, shows that

$$E[e^{-\vartheta T(u)}] = \exp(-u \vartheta), \ u \geq 0, \ \vartheta > 0,$$

where the function $\psi : [0, \infty) \to [0, \infty)$ is the Laplace exponent of $T$. Thus, there exists a unique pair $(k,d)$ of non-negative real numbers and a unique measure $\mu$ on $(0,\infty)$ satisfying

$$\int_{(0,\infty)} (1 \wedge x) \mu(\mathrm{d}x) < \infty, \ 	ext{such that} \ \psi(\vartheta) = k + d\vartheta + \int_{(0,\infty)} (1 - e^{-\vartheta x}) \mu(\mathrm{d}x), \ \vartheta > 0.$$

Conversely, any function $\psi$ can be expressed as (3.7), can be seen as the Laplace exponent of a subordinator (see, e.g. Bertoin, 1996). Here, we assume that $\psi(0) = 0$, and the drift $d = 0$. The constant $k$ is usually called the killing rate. By using condition (3.7), it is easy to check that $\lim_{t \to 0} t \mu([t, \infty)) = 0$ and $\int_{0}^{1} \mu([t, \infty)) \mathrm{d}t < \infty$. Furthermore, recall that a $C^\infty$ function $\psi : [0, \infty) \to [0, \infty)$ is called a Bernstein function if $(-1)^n D^n \psi \leq 0$ for every $n \in \mathbb{N}$. It is well known that a function $\psi : [0, \infty) \to [0, \infty)$ is Bernstein function if and only if it has representation given by (3.7). To this end, the results in Section 1.3 of James et al. (2008a) are relevant and of significant importance for this article. Specifically, the following definition is adopted.
Let $m > 0$ and let $G$ be a positive random variable such that $E[\ln^+(1/G)] < \infty$. We say that a positive random variable $G_m(G) := \int_0^m \frac{1}{F_G^{-1}(u/m)} d\gamma_u$ is a member of $(m,G)$ GGC (generalized gamma convolutions) if

$$E[e^{-d_{G_m}(G)}] = \exp\left(-m\int_0^\infty \left(1-e^{-x}\right)\frac{dx}{x} E[e^{-xG}]\right) = \exp\left(-mE\left[\ln\left(1+\frac{\lambda}{G}\right)\right]\right),$$

where $\{\gamma_t : t \geq 0\}$ is the usual gamma process, i.e., a subordinator such that:

$$P(\gamma_t \in dx) = p_t(x) dx, \quad p_t(x) = \frac{1}{\Gamma(t)} x^{t-1} e^{-x} I(x \in [0,\infty]), \quad t > 0 \text{ and } x \in \mathbb{R}.$$ 

Usually, the integral $G_m(G) := \int_0^m \frac{1}{F_G^{-1}(u/m)} d\gamma_u$ is called the Wiener-Gamma $(m,G)$ GGC. In addition, from Proposition 1.3 in James et al. (2008a), any Wiener-Gamma $(m,G)$ GGC $(m > 0)$ random variable satisfies

$$G_m(h) \overset{(law)}{=} G_m(G), \quad \text{and} \quad G = 1/h(U_m), \quad \text{for} \quad h(u) = 1/F_G^{-1}(u/m), \quad u \in [0,m].$$

where $U_m$ is uniformly distributed on $u \in [0,m]$.

To state the main theorem of this study, the $m$-Dirichlet means processes also play a significant role in the development of the theory of the random variable $J_m^+$. Thus, for any $m > 0$, the $m$-Dirichlet means process over $[0,m]$ is defined as

$$D^{(m)} = \{D^{(m)}_u : u \in [0,m]\} = \left\{\gamma_u : u \in [0,m]\right\}.$$ 

Note that $\gamma_m$ is independent of $D^{(m)}$. Further, as in (3.8), we introduce $(m,G)$ Dirichlet random variable as $D_m(G) := \int_0^m \frac{1}{F_G^{-1}(u/m)} dD^{(m)}_u$. For a survey of the main properties of $(m,G)$ Dirichlet
random variable see e.g., James et al. (2008a). It should be further highlighted here that gamma and Dirichlet processes have been the object of intensive research in recent years from both pure and applied viewpoints. In particular, they play an important role in representation theory of infinite dimensional groups, in mathematical finance and in mathematical biology. Thus, one of the main objects of this study is to relate the theory $J^+_\infty$ with both gamma and Dirichlet processes.

We shall now state the main theorem of this investigation.

**Theorem 3.1.** Let $\{X(t) : t \geq 0\}$ be as described in (3.6) with $\delta < 0$ and let $T = \{T(t) : t \geq 0\}$ be a subordinator, where its Laplace exponent satisfies (3.7) with $\psi(0) = 0$ and the drift $d = 0$.

Then, for $q = |\delta|/\sqrt{2}$, the following identities hold:

1. $E[\exp(-\lambda J^+_\infty)] = \exp\left(-\frac{1}{2} E_{\gamma_{1/2}} \left[ \ln \left( 1 + \frac{\lambda}{\psi(q^2/G_{1/2})} \right) \right]\right) = E\left[ 1 + \lambda D_{1/2} \left( \psi(q^2/G_{1/2}) \right)^{-1/2} \right], \lambda > 0$,

where $\psi : [1, \infty) \to [0, \infty)$ is a Borel function such that: $E\left[ \ln^+(1/\psi(q^2/G_{1/2})) \right]$.

2. If $E\left[ \ln^+(1/\psi(q^2/G_{1/2})) \right] < \infty$, then

   (i) $f_{J^+_\infty}(x) = \frac{\exp\left(E[\ln(\psi(q^2/G_{1/2})]/x^{1/2} \sqrt{\pi}\right)/2) E\left( \cos \left( 2 \sqrt{x} \Gamma_{1/2} \left( 1/\psi(q^2/G_{1/2}) \right) \right) \right)}{x^{1/2} \sqrt{\pi}} I(x \in (0, \infty))$,

   (ii) $f_{J^+_\infty}(x) = \frac{\exp\left(E[\ln(\psi(q^2/G_{1/2})]/x^{1/2} \sqrt{\pi}\right)/2) E\left[ \exp(-x D_{1/2} \left( 1/\psi(q^2/G_{1/2}) \right)) \right]}{x^{1/2} \sqrt{\pi}} I(x \in (0, \infty))$.

3. $J^+_\infty \overset{law}{=} \Gamma_{1/2} \left( \psi(q^2/G_{1/2}) \right) \overset{law}{=} \int_0^{1/2} \frac{d\gamma_n}{F^{-1}_{\psi(q^2/G_{1/2})}(2u)} = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \frac{1}{F^{-1}_{\psi(q^2/G_{1/2})}(2s_j)} (\gamma_{s_{j+1}} - \gamma_{s_j})$.

$G_{1/2}$ is beta distributed with parameters $(1/2, 1/2)$, $\Pi = \{s_0, s_1, \ldots, s_n\}$ is a partition of $(0, 1/2)$, such that $0 < s_0 \leq s_1 \leq \cdots \leq s_n < 1/2$ with norm $\|\Pi\| = \max_{0 \leq j \leq n} |s_{j+1} - s_j|$, and $\Gamma_t(G)$ is the Wiener-Gamma integral defined as (3.8).
The proof of the main theorem will be presented in the next section.

4. PROOFS OF THE MAIN RESULT

The proof of the main theorem is greatly enhanced by first establishing related results for the modified model (3.3) instead of the general case. The modified model (3.3) is now defined by setting the volatility, $\sigma_t = \sigma$, be constant and we call $\eta := \mu + \delta \sigma < 0$. The analytical details in developing the theory for $J^+_\infty$ under the model (3.3) will provide much needed insight for establishing the results in the general case of Theorem 3.1. The analysis begins by noting that under the linear (modified) Brownian motion model (3.3), we have

$$P(X(t) > 0) = P(\eta t + \sigma B(t) > 0) = \Phi(\sqrt{2qt^{1/2}}), \quad q = |\eta|/\sqrt{2} \sigma \text{ and } \eta < 0.$$ 

Then, the following proposition is in order.

**Proposition 4.1.** Let $X(t) = \mu t + \sigma B(t)$. Then, for $q = |\eta|/\sqrt{2} \sigma$, $\eta < 0$ and $\lambda \geq 0$, the following identities hold

1. $E[\exp(-\lambda q^2 J^+_\infty)] = \exp\left(-\frac{1}{2} \int_0^\infty \frac{du}{u} \left(1 - e^{-\lambda u}\right) \right) E\left[e^{-u/G_{1/2}}\right] = \frac{2}{1 + \sqrt{1 + \lambda}},$

2. $f_{q^2 J^+_\infty}(x) = \frac{1}{\sqrt{\pi}} x^{-1/2} \int_1^\infty e^{-xy} y^{-3/2} dy I(x \in [1, \infty))$

$$= \frac{1}{\sqrt{\pi}} x^{-1/2} \int_0^1 e^{-xy} y^{-1/2} dy I(x \in [0, \infty)), \text{ and}$$

3. $q^2 J^+_\infty \xrightarrow{(law)} \frac{1}{\Gamma_{1/2}(1/G_{1/2})} \frac{d\gamma^u}{2u} \xrightarrow{(law)} \int_0^{1/2} \sin^2(\pi u) d\gamma^u \xrightarrow{(law)} \gamma_{1/2} U.$
where $G_{\gamma/2}$ is a beta distributed random variable with parameters $(1/2,1/2)$. $U$ is uniformly distributed on $[0,1]$. $\gamma$ is a gamma distributed with parameters $(s,1)$ and $\Gamma_t(G)$ is the gamma-Wiener integral.

**Proof.** We begin by first expressing the survival function of the normal distribution in terms of the complement of the error function. Specifically, it can be seen that

\[
\Phi(\sqrt{2}qu^{1/2}) = \frac{\text{erfc}(qu^{1/2})}{2} \quad \text{where} \quad \text{erfc}(qu^{1/2}) = \frac{2q}{\pi} \int_0^\infty e^{-t_2 + q^2} \frac{dt}{t^2 + q^2} = e^{\gamma_u}, \quad q = |\eta|/\sqrt{2}\sigma.
\]

In light of (4.1), it follows that

\[
E[\exp(-\lambda J_\infty^+)] = \exp\left(-\int_0^{\infty} \frac{du}{u} (1 - e^{-\lambda u}) \Phi(\sqrt{2}qu^{1/2})\right)
\]

\[
= \exp\left(-\frac{q}{\pi} \int_0^{\infty} \frac{du}{u} (1 - e^{-\lambda u}) \int_0^{\infty} e^{-\frac{(t^2 + q^2)}{t^2 + q^2}} \frac{dt}{t^2 + q^2}\right)
\]

\[
= \exp\left(-\frac{q}{\pi} \int_0^{\infty} \frac{du}{u} (1 - e^{-\lambda u}) \int_0^{\infty} e^{-\frac{t^2 + q^2}{t^2 + q^2}} \frac{dt}{1 + t^2}\right)
\]

\[
= \exp\left(-\frac{1}{\pi} \int_0^{\infty} \frac{du}{u} \left(1 - e^{-\frac{\lambda u}{\pi^2}} \right) \int_0^{\infty} e^{-\frac{t^2}{1 + t^2}} \frac{dt}{1 + t^2}\right)
\]

\[
= \exp\left(-\frac{1}{2} \int_0^{\infty} \frac{du}{u} \left(1 - e^{-\frac{\lambda u}{\pi^2}} \right) \int_0^{\infty} e^{-\omega t} U(dt)\right).
\]

The last statement indicates that when $X(t)$, $t \geq 0$, is a linear Brownian motion with negative drift, $J_\infty^+$ is a self-decomposable random variable. Continuing the above calculations, the second integrand term in (4.2) may be expressed as

\[
\int_0^{\infty} e^{-\omega t} U(dt) = \frac{1}{\pi} \int_0^{\infty} \frac{e^{-\omega y} dy}{\sqrt{y-1}} = \int_0^{1} e^{-\omega y} f_{G_{\gamma/2}}(y) dy,
\]

\[
\text{SMTDA} \quad 222
\]
GAUSSIAN SUBORDINATED PROCESS AND ITS ARGUMENT SUPREMUM

where \( f_{G_{1/2}}(y) = \frac{1}{\pi} y^{-1/2} (1 - y)^{-1/2} \) is a probability density function of a beta random variable, \( G_{1/2} \), with parameters \((1/2,1/2)\). Substituting (4.3) into (4.2), we finally find that

\[
E[\exp(-\lambda J_\infty^+)] = \exp\left(-\frac{1}{2} \int_0^\infty \frac{du}{u} \left(1 - e^{-\lambda u}\right) E\left[e^{-u/G_{1/2}}\right]\right).
\]

Thus, applying Cifarelli and Melilli (2000) results (see also James et al., 2008a), it can be seen that

\[
E[\exp(-\lambda J_\infty^+)] = \frac{2}{1 + \sqrt{1 + \lambda q^{-2}}} = E\left[\exp\left(-\frac{\lambda}{q^2} \Gamma_{1/2}(1/G_{1/2})\right)\right].
\]

Equations (4.4) and (4.5) prove Proposition 4.1 part (1). Part (2) can be found in James et al. (eq. 198 and 226, 2008a). Part (3) can be found from the definition of \((m, G) \) GGC random variables or can be found in James et al. (e.g., eq. 197 or 199, 2008a).

To add more light to Proposition 4.1, an additional expression of the \( J_\infty^+ \) density is obtained by inverting the Laplace transform. Specifically, we let \( \mathcal{L}(f)\lambda = \int_0^\infty e^{-\lambda t} f(t)dt = \hat{f}(\lambda) \), \( \lambda \in [0, \infty) \), be the one–sided Laplace transform of a function \( f \) and let \( \mathcal{L}^{-1}(\hat{f})(t) = f(t), t \in [0, \infty) \) denote the corresponding inverse Laplace transform. Thus, for \( f \) continuous on \( t \in [0, \infty) \), and \( e^{-\lambda t} f(t), t \in [0, \infty) \) absolutely integrable, \( \hat{f} \) is an analytic function for \( \Re(\lambda) \in [0, \infty) \). It is then well known that \( \mathcal{L}^{-1}(af(\lambda) + b)(t) = af(t) + b, \) for any \( a, b \in \mathbb{R}, \) and \( \mathcal{L}^{-1}(\hat{f}(\lambda)/\lambda)(t) = \int_0^t f(u)du. \) Thus,

\[
2\left(1 + \left(1 + \frac{\lambda}{q^2}\right)^{1/2}\right)^{-1} = \lambda^{-1}(2q\left(\lambda + q^2\right)^{1/2} - q).
\]
Calling upon the above remarks the following corollary is then in order.

**COROLLARY 4.2.** Let \( \{X_t = \eta t + \sigma B_t : t \geq 0\} \) be a linear Brownian motion with \( \eta < 0 \) and \( \sigma > 0 \).

For \( q = |\eta|/\sqrt{2\sigma} \), the following holds:

\[
f_{q^2 J^{(o)}}(t) = \frac{2}{\sqrt{\pi}} \int_{-\infty}^{x} e^{-y^2} dy.
\]

**Proof.** From the above remarks, we have

\[
\mathscr{L}^{-1}\left(\frac{(\lambda + a)^{1/2}}{\lambda}\right)(t) = \frac{e^{-at}}{\sqrt{\pi t}} + a^{1/2} \text{erf}\left(\frac{\sqrt{at}}{\sqrt{2}}\right), \ a > 0 , \ \lambda \geq 0.
\]

Thus, for \( t > 0 \),

\[
\mathscr{L}^{-1}\left(2q \left(\frac{\lambda + q^2}{\lambda}\right)^{1/2} - 2 \frac{q^2}{\lambda}\right)(t) = 2q \frac{e^{-qt^2}}{\sqrt{\pi}} - 2q^2 + 2q^2 \text{erf}\left(\sqrt{q^2 t}\right)
\]

\[
= \frac{2}{\sqrt{\pi}} q \frac{e^{-qt^2}}{\sqrt{t}} - 2q^2 \text{erfc}\left(\sqrt{q^2 t}\right)
\]

\[
= \frac{4}{\sqrt{\pi}} q^2 \left\{ \frac{e^{-qt^2}}{2\sqrt{q^2 t}} - \int_{-\infty}^{\infty} e^{-y^2} dy \right\}
\]

\[
= \frac{2}{\sqrt{\pi}} q^2 \int_{q\sqrt{t}}^{\infty} e^{-y^2} dy,
\]

where \( \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-y^2} dy \geq 0 \). This completes the proof. \( \Box \)

**Proof of Theorem 3.1.** Adapting an approach similar to Proposition 4.1, it can be found that
GAUSSIAN SUBORDINATED PROCESS AND ITS ARGUMENT SUPREMUM

\[ E[\exp(-\lambda J^+_\infty)] = \exp \left( -\int_0^\infty \frac{du}{u} \left( 1 - e^{-\lambda u} \right) E\left[ \Phi\left( \sqrt{2}qT^{1/2}(u) \right) \right] \right) \]

\[ = \exp \left( -\frac{1}{2} \int_0^\infty \frac{du}{u} \left( 1 - e^{-\lambda u} \right) E\left[ e^{-\tau(u)^2/G_{1/2}^2} \right] \right), \]

where \( q = |\delta|/\sqrt{2}\sigma, \delta < 0. \)

Since \( T(t) \) is a subordinator, it follows that

\[ E\left[ e^{-\tau(u)^2/G_{1/2}^2} \right] = E\left[ e^{-u\psi(q^2/G_{1/2}^2)} \right], \quad u \geq 0. \]  

(4.7)

In light of equation (4.4) and the Frullani’s integral, it then implies that

\[ E[\exp(-\lambda J^+_\infty)] = \exp \left( -\frac{1}{2} E \left[ \ln \left( 1 + \frac{\lambda}{\psi(q^2/G_{1/2}^2)} \right) \right] \right). \]

(4.8)

Thus, equation (4.8) implies that \( J^+_\infty \overset{\text{(law)}}{=} \Gamma_{1/2}^\gamma \psi(q^2/G_{1/2}^2). \) It is known, from Bondesson (Th. 5.2.2, p.79, 1992), that \( \Gamma_{1/2}^\gamma \) is a member of GGC. Assuming that \( E[V^{-1/2}] < \infty \), James et al. (2008a) have shown that \( \Gamma_{1/2}^\gamma (h) = \gamma_{1/2} D_{1/2}(h) \) (i.e., the random variable \( V \) is expressed as \( \frac{1}{2} \)-Dirichlet integral), where \( h(u) = \int F_{\psi(q^2/G_{1/2}^2)}^{-1}(2u), \quad u \in [0,1/2], \) and \( D_{1/2}(h) = \int_0^{1/2} h(u) dD_u^{1/2} \) is the \( \frac{1}{2} \)-Dirichlet integral.

Calling upon the above statements, it is then easy to see that

\[ E[\exp(-\lambda J^+_\infty)] = E_{\gamma_{1/2}} E_{\Gamma_{1/2}^\gamma} \left[ \exp(-\lambda \Gamma_{1/2}^\gamma (\psi(q^2/G_{1/2}^2))) \right] = E\left[ \left[ \left[ 1 + \lambda \Gamma_{1/2}^\gamma (\psi(q^2/G_{1/2}^2)) \right]^{-1/2} \right] \right]. \]

(4.9)

Combining (4.7), (4.8) and (4.9) the proof of Theorem 3.1 part (1) is now established. Part (2) follows from the duality theorem found in James et al. (2008a). Specifically, using equation (102,
in James et al., 2008a), the density of the random variable \( J^+ = (\psi)^{\text{law}} \left( q^2 / G_{1/2} \right) \) can be expressed by

\[
f_{J^+}(x) = \exp\left( E\left( \ln(\psi(q^2 / G_{1/2})) \right) \right) / 2
\]

(4.10)

\[
E \left[ \left( \frac{\Gamma_{1/2}(1/\psi(q^2 / G_{1/2}))}{x} \right)^{1/4} \cdot J_{-1/2} \left( 2 \sqrt{x \Gamma_{1/2}(1/\psi(q^2 / G_{1/2}))} \right) \right] I(x \in (0, \infty)),
\]

where \( J_\nu \) denotes the Bessel function with index \( \nu \):

\[
J_\nu(z) = \sum_{j=0}^{\infty} (-1)^j \frac{1}{j! \Gamma(j+\nu+1)} \left( \frac{z}{2} \right)^{\nu+2j}, \quad |z| < \infty \text{ and } |\arg z| < \pi.
\]

Upon substituting \( J_{-1/2}(x) = \sqrt{2 \pi x \cos x} \) into (4.10), the proof of Theorem 3.1 (2) is now in order.

Finally, part (3) follows from the definition of the \((m, G)\) GGC \((m > 0)\) Wiener-Gamma integral.

**REMARK 4.3.** The formulae given in Theorem 3.1 may be expressed in several other ways depending on various expressions of the \(1/2\)-Dirichlet integral (see James et al. 2008a). Specifically, it can be seen that

\[
\gamma_{1/2}^D \gamma_{1/2}^D (1/G)^{(law)} = \gamma_{1} \gamma_{1/2}^D (1/G)^{(law)} = \gamma_{1} D_{1/2}(1/G Y_{1/2}),
\]

where, again, \( \gamma_1 \) and \( G_{1/2} \) are independent, \( Y_{1/2} \) and \( G \) are independent and \( Y_{1/2} \) is a Bernoulli random variable with probability \( 1/2 \).

**REMARK 4.4.** Further, the density of \( D_1(1/G Y_{1/2}) \) (see James et al. 2008a) may be expressed as

\[
f_{D_1(1/G Y_{1/2})}(x) = \frac{\sin(\pi G_x(x^{-1})/2)}{\pi^{1/2}} \exp(-E[\ln(x - G^{-1}) I(G \neq x^{-1})]/2).
\]
Remarks 4.3 and 4.4 apply to the second expression of the Laplace transform for \( J^+ \) in Theorem 3.1. The same remarks also apply to the various examples considered in the next section.

5. EXAMPLES

We shall now illustrate the utility of the main theorem by computing various parts of Theorem 3.1 for some of the familiar subordinators, as examples. To present the stable one, we begin by introducing the \( \alpha \)-stable Lévy motion.

A stochastic process \( T = \{ T(t) : t \in \mathbb{R}_+ \} \) is an (standard) \( \alpha \)-stable Lévy motion if

1. \( T(0) = 0 \) a.s.
2. \( T \) has independent increments.
3. \( T(t) - T(s) \sim S_{\alpha}((t - s)^{1/\alpha}, \beta, 0) \), for any \( 0 \leq s < t \) and for \( \alpha \in (0, 2] \).

In order that \( T \) be a subordinator, we further require that \( \alpha \in (0, 1) \) and \( T \) to be totally skewed to the right, i.e., \( \beta = 1 \). Thus, the example for the stable case is now formulated as follows.

**EXAMPLE 5.1.** Let \( X(t) = \delta T(t) + T^{1/2}(t)N \), \( \delta < 0 \) and let \( T = \{ T(t) : t \geq 0 \} \) an \( \alpha \)-stable Lévy motion which is totally skewed to the right and \( \alpha \in (0, 1) \). Then, for \( q = |\delta|/\sqrt{2} \), \( \delta < 0 \) and \( c = \cos(\pi \alpha/2)/q^{2\alpha} \), the following identities hold:

1. (i.) \( E[\exp(-\lambda J^+) = \exp\left(-\frac{1}{2}E_{G_{1/2}}\left[\ln\left(1 + \lambda cG_{1/2}^\alpha\right)\right]\right) = E\left[\left(1 + \lambda D_{1/2}(1/cG_{1/2}^\alpha)\right)^{-1/2}\right] \), \( \lambda \geq 0 \),

2. \( f_{J^+}(x) = \frac{e^{-E[\ln(cG_{1/2}^\alpha)]/2}}{x^{3/2} \sqrt{\pi}} E\left[\cos\left(2\sqrt{xT_{1/2}(cG_{1/2}^\alpha)}\right)\right] I(x \in (0, \infty)) \)
where $G_{1/2}$ is a beta distributed random variable with parameters $(1/2,1/2)$, $\Pi = \{s_0,s_1,\cdots,s_n\}$ is a partition of $(0,1)$, such that $0 < s_0 \leq s_1 \leq \cdots \leq s_n < 1/2$ and $\|\Pi\| = \max_{0 \leq j \leq n} |s_{j+1} - s_j|$, $\gamma_s$ is a gamma distributed with parameters $(s,1)$ random variable and $\Gamma_t$ is the gamma-Wiener integral.

The only statement that requires some attention is part 1(ii). The rest are follow-ups from the proof of the main result. For clarity, it is known that if $T(t)-T(s) \sim S_{\alpha}(t-s)^{1/\alpha},(1,0)$, for any $0 \leq s < t$ and for $\alpha \in (0,1)$, then it follows that

$$E\left[ e^{-T(u)/G_{1/2}} \right] = E\left[ e^{-\psi(\alpha \pi/2)/G_{1/2}} \right].$$

In light of equation (5.1) and the Frullani’s integral, it then implies that

$$E[\exp(-\lambda J^+_x)] = \exp\left(-\frac{1}{2} E\left[ \ln(1 + \lambda c G_{1/2}^\alpha) \right] \right).$$

The rest of the proof is then clear.

To add some insight on the random variable $J^+_x$ in part 3, we note that

$$F_{1/G_{1/2}}(x) = P(1/G_{1/2} \leq x) = P(G_{1/2} \geq x^{-1/\alpha}) = 1 - F_{G_{1/2}}(x^{-1/\alpha}), \text{ for } x \in (0,1].$$

Thus, it can be easily seen that $F^{-1}_{1/G_{1/2}}\left(l - F_{G_{1/2}}(x^{-1/\alpha})\right) = x, x \in (0,1]$. Using part 3, it also follows that

$$2s = 1 - F_{G_{1/2}}(x^{-1/\alpha}) \text{ or } \left(F^{-1}_{G_{1/2}}(1 - 2s)\right)^{-\alpha} = x, \text{ for } x \in (0,1] \text{ and } s \in (0,1/2].$$

**Remark 5.2.** It is important to justify that condition $E_{G_{1/2}}\left[\ln(\psi(1/G_{1/2}))\right] < \infty$ is valid when a stable subordinator is considered. The above condition is then equivalent to $E_{G_{1/2}}\left[\ln(G_{1/2})\right] < \infty$. 

24
which is dominated by $E_{G_{q,2}^{-1}}[G_{1/2}^{-1}] < \infty$, such that $\frac{1}{2} - \varepsilon > 0$. This in turn shows that condition $E_{G_{q,2}^{-1}} \left[ \ln \left( G_{1/2} \right) \right] < \infty$ does hold.

**Example 5.3.** We now assume that $\{T(t) : t \geq 0\}$ is a standard gamma process. Obviously, we have that $\{T(t) \equiv \gamma_t : t \geq 0\}$ is a subordinator without drift with Lévy measure $\frac{dx}{x} e^{-x} I(x > 0)$.

Thus, the Lévy-Khintchine representation is expressed as

$$E[e^{-\lambda T_t}] = \exp(-u \ln(1 + \theta)) = (1 + \theta)^{-u}, \quad u \geq 0.$$  

Thus, when $\{T(t) : t \geq 0\}$ is a gamma subordinator, the following identities can be derived:

1. $E[\exp(-\lambda J^*_\infty)] = \exp \left( -\frac{1}{2} E \left[ \frac{\lambda}{\ln(1 + q^2 G^{-1}_{1/2})} \right] \right) = E \left[ 1 + \lambda D_{1/2} \left( \ln \left( 1 + q^2 G^{-1}_{1/2} \right) \right) \right]$, $\lambda \geq 0$,

2. $f_{J^*_\infty}(x) = \frac{\exp \left( E \left[ \ln \left( 1 + q^2 G^{-1}_{1/2} \right) \right] \right)}{x^{1/2} \sqrt{\pi}} \exp \left( -\frac{1}{2} E \left[ \cos \left( 2 \sqrt{x T_{1/2}^{-1} \left( \ln \left( 1 + q^2 G^{-1}_{1/2} \right) \right)} \right) \right] I(x \in (0, \infty))$, and

3. $J^*_\infty \overset{\text{(law)}}{=} \Gamma_{1/2} \left( \ln \left( 1 + q^2 G^{-1}_{1/2} \right) \right) \overset{\text{(law)}}{=} \int_0^{1/2} \frac{d\gamma_u}{F_{\ln(1+q^2 G^{-1}_{1/2})}^{-1}(2u)} = \lim_{\theta \downarrow 0} \frac{1}{\theta} \sum_{j=0}^{n-1} \frac{1}{F_{\ln(1+q^2 G^{-1}_{1/2})}^{-1}(2s_j)} (\gamma_{s+1} - \gamma_{s})$.

**Example 5.4.** Here, we let $T(t), t \geq 0$, be a simple Poisson process with parameter $\theta$, $\theta \geq 0$. Note that $T$ takes values on $\mathbb{N}$ (natural numbers), it is an additive process and its Laplace transform can be expressed by $E[e^{-\lambda T(t)}] = \exp(-t(1 - e^{-\theta})) \quad \lambda \geq 0$.

In light of (4.8) and (4.9) and the remarks made above, it then follows that

$$E[\exp(-\lambda J^*_\infty)] = \exp \left( -\frac{1}{2} \int_0^\infty \frac{du}{u} \left( 1 - e^{-\lambda u} \right) E[e^{-\lambda T_{u q^2 / G_{1/2}}} \right]$$

$$= \exp \left( -\frac{1}{2} E \left[ \ln \left( 1 + \frac{\lambda}{\theta(1 - e^{-\lambda q^2 / G_{1/2}})} \right) \right] \right), \lambda \geq 0 \text{ and } \theta \geq 0.$$
Thus, \( J^+_{\infty} \overset{\text{law}}{=} \Gamma_{3/2} \left( \theta \left( 1 - e^{-q^2/G_0} \right) \right) \), which shows that

\[
E[\exp(-\lambda J^+_{\infty})] = E\left[ \left( 1 + \frac{1}{\theta} \left( 1 - e^{-q^2/G_0} \right) \right)^{-1/2} \right].
\]

In determining the density, we just imitate similar arguments as earlier and conclude that the density's expression is given by

\[
f_{J^+_{\infty}}(x) = \exp\left( \frac{E[\ln(\theta(1-e^{-q^2/G_0}))]/2}{x^{1/2} \sqrt{\pi}} \right) \cos\left( 2 \sqrt{\pi} J_{1/2}^{1/2} \left( 1 - e^{-q^2/G_0} \right) \right) I(x \in (0, \infty)).
\]

Finally, the stochastic representation of the ultimum argument maximum for \( J^+_{\infty} \) is:

\[
J^+_{\infty} \overset{\text{law}}{=} \Gamma_{1/2} \left( \theta \left( 1 - e^{-q^2/G_0} \right) \right) \overset{\text{law}}{=} \int_0^{1/2} \frac{\text{d}y_u}{F^{-1}{\theta \left( 1 - e^{-q^2/G_0} \right)}}(2u) = \lim_{n \to \infty} \sum_{j=0}^{n-1} \frac{1}{F^{-1}{\theta \left( 1 - e^{-q^2/G_0} \right)}}(2s_j)(\gamma_{s_{j+1}} - \gamma_s).
\]

**Remark 5.5.** It should be pointed out that the second identity in part 1 of the Examples 5.1, 5.3 and 5.4, and Theorem 3.1 can be alternatively shown using the Dirichlet Process presentation. Specifically, the process can be expressed as a normalized gamma process and then the identity follows from the Markov-Krein identity for distributions of Dirichlet means processes (see, e.g., Vershik, Yor and Tsilevich, 2004, or Tsilevich and Vershik, 1999).

**Remark 5.6.** The restriction on the parameter \( \alpha \in (0,1) \) in Example 5.1 is quite natural. It can be easily seen that when \( \alpha \in (1,2) \), the quantity \( c = \cos(\pi \alpha/2) / q^{2\alpha} \) becomes negative, but more importantly the resulting process is not anymore a subordinator.
6. NUMERICAL RESULTS

In this section, we illustrate the distribution of $J_{\alpha}^+$ for the three example cases of Section 5, where the process $T(t)$ follows: (i) an $\alpha$-stable subordinator with choices of $\alpha = 0.25, 0.5, 0.75$, (ii) a gamma process with parameter 1, and (iii) a Poison process with parameter $\theta = 1$. In addition in case (iv), we let the process $X(t)$ be simply a linear Brownian motion with drift $\eta = -0.1$ (see model (3.3) under $\eta := \mu + \delta \sigma < 0$); in this case $T(t) \equiv t$. Note that the simple linear Brownian motion case corresponds to $\alpha$-stable subordinator with $\alpha = 1$, with the results for this case appearing in Proposition 4.1. The goal is to compute the theoretical density function of $J_{\alpha}^+$ by numerically inverting the Laplace transform that appears in Part (1) in all three examples as well as computing it empirically through the representation in Part (3) with the help of 500,000 simulations in each case. In carrying out these computation we consider the drift coefficient to be $\delta = -0.1$. The graphs of the density for $J_{\alpha}^+$ are represented in Figure 6.1(a)-(c) for case (i) and Figure 6.1(d) for case (iv); Figure 6.2(a) for case (ii) and Figure 6.2(b) for case (iii). It may be noted that there is a close match between the theoretical density and the empirical density in all cases.
Figure 6.1. Theoretical density (continuous curve) and empirical density (histogram) of $J_{\alpha}^+$ for $\alpha$-stable subordinator in case (i) with (a)-(c) representing $\alpha = 0.25, 0.5, 0.75$, respectively, and (d) representing case (iv) for the linear Brownian motion with $\alpha = 1$.

Figure 6.2. Theoretical density (continuous curve) and empirical density (histogram) of $J_{\alpha}^+$ with (a) representing gamma process in case (ii), and (b) representing the Poisson process in case (iii).
7. CONCLUDING REMARKS

In this article, we consider the problem of deriving the distribution of ultimum argument supremum of Gaussian subordinated processes with negative drift. The limiting distributions we derive are for the non-normalized ultimum argument supremum, whereas some of the classical results such as the reflection principle are derived under proper normalization. The results we derive show that the ultimum argument supremum can be represented by the recently well-studied Wiener-Gamma integrals with a constant parameter 0.5. The motivation for considering the Gaussian subordinated processes stems from the fact that such processes have been the central focus of research and application in the recent financial literature. For example, the normal inverse Gaussian process is a member of the Gaussian subordinated process. Numerical computations of the density for the ultimum argument supremum show close agreement between the theoretical results derived and the corresponding empirical approximations.

REFERENCES


GAUSSIAN SUBORDINATED PROCESS AND ITS ARGUMENT SUPREMUM


Poisson Changepoint Data and Goodness of Fit

James M Freeman & Shuming Chen

Manchester Business School, University of Manchester, Booth Street West, Manchester, M15 6PB, UK
Email: jim.freeman@mbs.ac.uk
chenshuming828@126.com

Abstract: In the paper we present details of a statistical procedure for detecting an unknown change-point for a sequence of Poisson variables. The new methodology is based on a goodness of fit formulation. It has been tested on a variety of simulated and actual datasets, including a number, well-known from the literature.

The procedure has been found to be particularly appropriate to problems where either an abrupt change or a cumulative change in the value of a Poisson parameter has occurred after an unknown point.

Keywords: Goodness of Fit, Poisson Process, Changepoint

1 Introduction

A sequence of independent Poisson variables is subject to a change in distribution after an unknown point. Formally, we can describe this situation as follows: $X_1, X_2, X_3, ..., X_k$ are independent random variables that occur in the $i$th fixed time interval, such that, for a value of $\tau$,

$$
X_i \begin{cases}
\text{is distributed as } & P(\lambda_0) \\
& (i = 1, 2, ..., \tau) \\
\text{or } & P(\lambda_1) \\
& (i = \tau+1, \tau+2, ..., k)
\end{cases}
$$

(1)

Previous work, with this type of model, has been directed toward (i) estimating the change-point, $\tau$ and (ii) on testing the hypothesis that no change in distribution has occurred.

Most analytical approaches, developed for dealing with Poisson change-point data, assume the $\lambda_0$ and $\lambda_1$ parameters, like $\tau$, to be unknown.

The problem has been analysed from a variety of perspectives: Asymptotic methods employed by Akman and Raftery (1992) compare with maximum likelihood alternatives by Henderson and Matthews (1993) and West and Ogden (1994). In contrast, Raftery and Akman (1986) employ Bayesian methodology. Alternative techniques, such as those of Worsley (1986), Loader (1992) and Perry (2006) provide further scope for analysis.
The paper introduces a new procedure for analysing Poisson change point data. The procedure is more general than many of the techniques developed in this area: it is not only capable of monitoring problems involving a single change in parameter level but also those where the change in parameter level has been cumulative after some unknown point. The technique, based on an unusual 'goodness of fit' argument is illustrated on a number of relevant data sets mostly from the literature.

2 Analysis

Referring to model (1), we distinguish between the rival sets of assumptions:

\[ H_1: \lambda_0 \text{ and } \lambda_1 \text{ are fixed with } \lambda_0 \neq \lambda_1 \]

\[ H_{1}': \lambda_0 \text{ is fixed and } \lambda_1 \text{ is a linear function of pre-chosen scores } \mathbf{s}_i \text{ which we write as } \lambda_1 = \lambda_0 + \beta (\mathbf{s}_i - \bar{s}) \]

where \( \beta \) is fixed

There are several methods with which the score \( \mathbf{s}_i \) could be chosen. Williams (1988) and Freeman (2009) show a number of available options.

Under the null hypothesis \( H_0 \) we assume no change in \( \lambda_0 \) 2 has taken place and therefore write

\[ H_0: \tau = k \quad \text{or} \quad H_0: \lambda_i = \lambda_0 3 \]

When either \( H_1 \) or \( H' \) holds, it is often found from a plot of \( X_i \) against \( i \) - that there appears to be a linear association between the two variables. Such an association would normally be investigated using Poisson regression. Nowadays, Poisson regression is mostly carried out using generalised linear modelling (GLM) procedures (Cameron and Trivedi, 1998) but historically – for preliminary data analysis at least – was often carried out using OLS methods where the \( X_i \) counts were first subjected to the log transform. This had the advantage, not only of eliminating skewness in the original data but also of inducing constant variance in the same. In the case of zero counts, because \( \ln(0) \) is not defined, an additional refinement was to add a constant of say 1 or 0.5 to the \( X_i \) before making the log conversion. In the regression-based analysis that follows, values of the dependent variable are therefore taken to be \( x_i = \ln (1+X_i) \).
For OLS regression the R square statistic is commonly adopted for representing the strength of a regression relationship. It can be shown that the latter statistic can be adapted to suit the specific circumstances of the change-point problem. Let

\[ R^2_c = \frac{S_{2t} - S_{2t}}{S_{2t}} \quad (t = 2, 3, \ldots, k-2) \quad (2) \]

where \( S_{2t} \) and \( S_{2t} \) correspond with the sums of squares from the regression of the \( x_i \) ratios on their index \( i \) for the first \( t \) and last \( (k-t) \) observations respectively and \( T_{1t} \) and \( T_{2t} \) are the corresponding corrected sums of squares on \( x_c \).

In particular:

\[ S_{2t} = \frac{\left( \sum_{i=1}^{2t} (x_i - \bar{x}) (y_i - \bar{y}) \right)^2}{\sum_{i=1}^{2t} (y_i - \bar{y})^2} \]

\[ S_{2t} = \frac{\left( \sum_{i=t+1}^{2t} (x_i - \bar{x}) (y_i - \bar{y}) \right)^2}{\sum_{i=t+1}^{2t} (y_i - \bar{y})^2} \quad (3) \]

where \( R_{1t} = \frac{1}{t} \sum_{i=1}^{t} x_i \) and \( R_{2t} = \frac{1}{(k-t)} \sum_{i=t+1}^{k} x_i \).

The formulas of \( T_{1t} \) and \( T_{2t} \) are as follows:

\[ T_{1t} = \sum_{i=1}^{t} (x_i - R_{1t})^2 \]

\[ T_{2t} = \sum_{i=t+1}^{k} (x_i - R_{2t})^2 \quad (4) \]

The \( R^2 \) statistic is analogous to that used by Freeman (1986) in an analysis of normal change-point data. Straightforward application of Freeman's methodology to model (1) confirms the following estimation procedure to be appropriate:

Under hypothesis \( H_1 \), estimate the change-point \( \tau \) as the value of \( t \) at which \( R^2 \) is minimised.
Under H_1, estimate \( \tau \) as the value of \( t \) at which \( R_t^2 \) is maximised.

Under \( H_0 \), the latter distribution can be shown to tend asymptotically to:

\[
B(\Gamma((k-4)/2)) \frac{((k-2)/2)}{((k-4)/2)}(1 - R_t^{2(k-6)/2})
\]  

(5)

Though the \( R_t^2 \) variates themselves are highly correlated, a Type 1 extreme value distribution can be used as an approximation to the distribution of the maximum value of \( R_t^2 \). See Freeman (1986) for details and corresponding critical values. By default, the distribution of the minimum value of \( R_t^2 \) can also be determined.

### 3. Applications

#### 3.1 Elliott and Shope’s Crash Rate data

Twenty eight observations on 16 year-old drivers’ crash rates in Michigan were reported by Elliott and Shope in 2003. The data were broken down by a) all crashes, b) night time and c) single vehicle crashes for the quarters January-March 1994 through to October-December 2000.
Figure 1, shows the $R^2_t$ plot for the all crashes series, the minimum value of $R^2_t (= 0.0044)$ occurring at the $t=14$ which corresponds with the second quarter of 1997. This is the time the new graduated drivers’ licensing program went into effect. Using the critical values referred to in section 2, it can be shown this result for $\min R^2_t$ is significant at the 5% level.

Of interest, for the same data, Elliot and Shope using Bayesian methods, estimated exactly the same changepoint and this was similarly found to be significant.

3.2 Jarrett’s Mining Disasters data

The second data set concerns the number of annual coal mining disasters – due to coal dust explosions - involving 10 or more fatalities from 1851 to 1962 (Jarrett 1979). This data was first presented by Maguire, Pearson, and Wynn (1952), but corrected and expanded by Jarrett in 1979.

For the purpose of the $R^2_t$ analysis the data was reduced in size by totalling the values four years at a time to make the calculations more manageable. Based on this summarisation it was found that the minimum value of $R^2_t (= 0.1049)$ occurred at $t = 10$, this point corresponding to the period 1887 - 1890.

This result compared with the estimates obtained by previous analysts (West and Ogden, 1994) as follows:
Proceedings, 2nd Stochastic Modeling Techniques and Data Analysis International Conference
5 - 8 June 2012 Chania Crete Greece

<table>
<thead>
<tr>
<th>Method</th>
<th>Change-point $T$</th>
<th>Left limit</th>
<th>Right limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loader (1992)</td>
<td>Mar 10, 1890</td>
<td>Length 8.08 years</td>
<td></td>
</tr>
<tr>
<td>Akman and Raftery (1992)</td>
<td>Mar 10, 1890</td>
<td>Oct 6, 1886</td>
<td>Dec 17, 1898</td>
</tr>
<tr>
<td>Worsley (1986)</td>
<td>1890</td>
<td>1884</td>
<td>1895</td>
</tr>
</tbody>
</table>

Table 1: Previous point and interval estimates of the change-point

Mirroring our own motivation, Adrian Raftery in his analysis had been concerned to identify if the reduction in the number of disasters had been gradual or abrupt. Corroborating his inference that it had been abrupt, one of his historian friends discovered that in 1889, the miners had formed the militant Miners' Federation (which later became the National Union of Mine Workers). Safety was their number one issue. Almost overnight, coal mines became safer (McGrayne, 2010).

3.3 Simulated data
Data were simulated by the authors as follows: assuming a change-point to hold at $T = 8$, $\lambda_i$ 17 18 was taken to be

$$\begin{align*}
\lambda_i & = \begin{cases} 
7 & (i = 1,2,...,8) \\
7 + 0.2 \, i & (i = 9,10,...,38)
\end{cases}
\end{align*}$$

The data are shown in Table 2 below:

<table>
<thead>
<tr>
<th>Interval (i)</th>
<th>Occurrences ($X_i$)</th>
<th>Interval (i)</th>
<th>Occurrences ($X_i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>23</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>25</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>26</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>28</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>29</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>30</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>31</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>32</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>9</td>
<td>33</td>
<td>13</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>34</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>11</td>
<td>35</td>
<td>10</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>36</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>37</td>
<td>10</td>
</tr>
<tr>
<td>19</td>
<td>9</td>
<td>38</td>
<td>6</td>
</tr>
</tbody>
</table>
4 Conclusions

A novel approach to the identification and testing of an unknown change point for a series of Poisson variates has been introduced. The approach has been demonstrated to have the particular advantage of distinguishing between situations where a Poisson parameter undergoes an abrupt as opposed to a cumulative value change. Results obtained using the procedure, have been found to compare well with established alternatives across a number of archetypal data sets. However significance testing is presently a weakness. No formal or empirical percentage points for this particular adaptation exist at the moment and those used earlier based on the normal distribution have yet to be properly validated for their applicability. Notwithstanding the latter caveats, the technique appears to offer considerable promise, and we look forward to a time when it might be possible for, the OLS basis of the modelling to be brought more in line with contemporary, better favoured GLM usage.

References

Estimate Effects of a Graduated Driver’s Licensing Program”. Journal of Data
Science 1, 43-63.
Statistician 35 3, 335-344.
5. Freeman, J.M. (2009) “Inference for Binomial Change-point Data”. In: Skiadas
Biometrika 66 1, 191-193.
Statistics 20 3, 1391-1411.
between Industrial Accidents”. Biometrika, 38, 168-180.
rule cracked the enigma code, hunted down Russian submarines, and emerged
triumphant from two centuries of controversy. Yale University Press. New Haven
and London.
point of a Poisson Rate Parameter with a Linear Trend Disturbance”, Quality and
Reliability Engineering International 22. 4, 371-384.
with A Change-Point”. Biometrika 73 1, 85-89.
Point in a Poisson Process”. Department of Statistics, University of South
Carolina Columbia.
Sequence of Exponential Family Random Variables”. Biometrika 73 1, 91-104.
Optimizing the Resources of a Healthcare Department using a Genetic Algorithm

Jennifer Gillespie1,2, Sally McClean1,2, Francis FitzGibbons2, Bryan Scotney1, Frank Dobbs3 and Brian Meenan2

1School of Computing and Information Engineering, University of Ulster, Northern Ireland
Email: gillespie-j@email.ulster.a.uk, {si.mcclean, bw.scotney}@ulster.ac.uk
2 Multidisciplinary Assessment of Technology Centre for Healthcare Programme, School of Engineering, University of Ulster, Northern Ireland
Email: {f.fitzgibbon, bj.meenan}@ulster.ac.uk
3Faculty of Life and Health Sciences, University of Ulster, Northern Ireland, UK
Email: f.dobbs@ulster.ac.uk

Abstract: Musculoskeletal disorders affect over 100 million people in Europe and with our population ageing this number is expected to rise which will put increased pressure on orthopaedic departments. The orthopaedic Integrated Clinical Assessment and Treatment Service (ICATS) was implemented in the United Kingdom by the Department of Health (DOH) to reduce the bottleneck within the department. However, due to unanticipated arrivals the DOH are aware that queues have started to build up again. In previous work we implemented a resource allocation methodology, based on Classic Queueing Theory, to find the minimum number of resources required for the system to reach steady state. In this paper we have extended our approach to develop a methodology which finds the best way to distribute a fixed number of resources through the healthcare department to minimize the overall waiting time. Using an Exhaustive Search Algorithm to find an optimal solution would take a lot of computational power and time therefore, to achieve our desired result we will use a Genetic algorithm. A Genetic algorithm is a heuristic search based on the process of natural evolution. The results show that although the Genetic Algorithm does not find the global optimum solution it finds a very close approximation and would be beneficial to use when there are a large number of resources to be distributed between many stages of a department as is commonly the case.

Keywords: Stochastic modeling, Orthopaedic ICATS, Genetic Algorithm, Queueing Theory

1 Introduction

It is estimated that musculoskeletal disorders affect more than 100 million people in Europe and 60% of patients on long term sickness in the United Kingdom (UK) state musculoskeletal problems as their reasons (Arthritis Foundation, 2009; Department of Health, 2006). Our population is ageing therefore these figures are expected to rise which will have a severe impact on the efficiency of orthopaedic departments.

The orthopaedic Integrated Clinical Assessment and Treatment Service (ICATS) was implemented in England, Scotland and Wales in 2005 to reduce the bottleneck that was created by patients waiting for an appointment with an orthopaedic surgeon. This ICATS system was implemented in Northern Ireland in
2007 and as it is the region of the UK with the highest waiting lists we will use data from the Southern Health and Social Care Board in this paper (Department of Health, 2006). The ICATS system has now been employed for over 5 years in Great Britain and over 3 years in Northern Ireland and in this time it has evolved. The orthopaedic process now begins with a patient referral being assessed by a healthcare professional (Physiotherapist, Podiatrist or General Practitioner with a special interest in Orthopaedics) in Paper Triage where they decide on the most appropriate treatment pathway (Southern Health and Social Care Trust, 2010). The pathways include going for a Face to Face (F2F) assessment with an ICATS health professional (55%), having diagnostic tests (7%), being sent to another department for treatment (10%), being referred back to their General Practitioner with management advice (3%) or having surgery (19%). It is also known that 6% of orthopaedic referrals do not attend their appointment. If a patient is sent to F2F assessment then 60% of those patients will be sent to Treatment and Review (T&R) appointments. On average patients receive three T&R sessions before being discharged, 20% of patients referred to F2F assessment will be sent to their GP with management advice and 20% will be sent to an orthopaedic consultant to have surgery 1 (Southern Health and Social Care Trust, 2010). A visual representation of the process along with the corresponding percentages can be seen in Figure. This new process tries to ensure that the patients’ waiting time is reduced, that they see the correct healthcare professional first time and that the consultant’s time is better spent in surgery with patients who require their expertise then in outpatient appointments (Rymaszewski et al. 2005).

![ICATS Process Diagram](image)

Figure 1. Orthopaedic ICATS Process

In 2010 the Southern Health and Social Care Board produced a document (Southern Health and Social Care, 2010) which stated that due to the number of
unanticipated arrivals to Orthopaedic ICATS the queues within the system are building up and the resources are unable to cope with the demand. In previous work (Gillespie, et al. 2011) we have developed a resource allocation methodology which finds the minimum number of resources required within the department for the system to reach steady state. In this paper we will extend this work to develop a methodology which finds the best way to distribute a fixed number of resources through a healthcare department to minimize the overall waiting time. To achieve this we use a genetic algorithm (GA) which is a search heuristic which mimics the process of natural evolution.

2 Optimizing Health Care Departments using a GA

In many situations it can be difficult to find an exact solution to a problem. There could be many reasons for this including the complexity of the objective function and the computational power required. In these cases it may be sufficient to use a technique, such as the Genetic Algorithm, to find a local optimum or to be at least close to a local or global optimum.

The Genetic Algorithm (GA) belongs to a larger class of evolutionary algorithms and can be used as a search heuristic which mimics the process of natural evolution. The algorithms can be used to generate useful solutions to optimisation and search problems within many sectors. The GA differs from other conventional methods as it does not search from a single point but operates on a whole population of points. This improves the chance of the algorithm reaching the global optimum and reduces the risk of becoming trapped in a local optimum. As the normal genetic algorithm does not use any auxiliary information about the objective function values, such as the derivatives, it can be applied to any kind of continuous or discrete optimization problem. This means that the GA is versatile and can be used within many different application areas to optimize a function by finding the best solutions (Holland, 1975).

Given a specific problem to solve a user is require to provide the GA with a set of potential solutions which can be quantitatively evaluated within a metric called the objective function. This can be done for a set of solutions which the user already knows work but in most cases these are done at random. The GA evaluates each solution according to a fitness function but in many cases, when the candidates are randomly selected from a pool, this will not work and infeasible solutions will be deleted. In some cases the solution is promising and these are kept and allowed to reproduce, therefore producing multiple ‘offspring’ (new solutions). However, these solutions may not be optimum and the GA allows random changes to be introduced in the copying process. These ‘offspring’ are then the potential solutions for the next generation. This process is repeated with the expectation that the average fitness of the function will increase with each generation until the optimal solution is found. There are many different selection methods which could be used to select the individuals to be copied over to the next generation, these include the Elitist method, Roulette wheel selection, Scaling selection and Rank selection (Marczyk, 2004).
In this paper we have developed a methodology for finding the best way to distribute a fixed number of resources to a healthcare facility by applying a genetic algorithm. Let us consider a healthcare department with \( i \) distinct consecutive stages, where \( i = 1, \ldots, n \). Let the arrival rate and service rate, \( \lambda_i \) and \( \mu_i \) respectively, follow an exponential distribution, \( c_i \) be the number of resources available to each stage and let there be an infinite waiting room/list for each stage. This can be seen as a network of infinite queues \( \text{M/M/}c_i \) and can be analysed using queueing theory. If we have a fixed number of resources, \( C \), available to the department we can use the GA to find the best way to distribute the resources in order to minimize the waiting time at each stage. The overall waiting time for such a network of infinite queues can be found using Little’s formula from classic queueing theory (Koizumi, et al. 2005).

The objective for the GA is to minimize:

\[
\min f(c) = \sum_{i=1}^{n} \left[ \frac{e^{-\lambda_i c_i} \sum_{n=0}^{c_i-1} \frac{\lambda_i^n c_i!}{n!} (1-\rho_i) c_i!}{\rho_i c_i!} \right]^{-1} \lambda_i
\]

where \( c=(c_1, \ldots, c_n) \), \( n_i \) (\( < c_i \)) is the number of occupied resources up the maximum number of resources \( c_i \), \( \omega = \lambda / \mu_i \) and \( \rho = \omega / c_i \) in a steady state system. For more information on this equation see Koizumi, et al. (2005).

The optimization problem is then solved subject to the constraints:

\[
Ac \leq b, \quad \sum_{i=1}^{n} c_i = C, \quad c_i > 0 \text{ for } i=1, \ldots, n
\]

where \( A \) is an \( (n \times n) \) diagonal matrix consisting of the maximum utilization value, \( \alpha \), for each diagonal element. The utilization value states how busy the system is allowed to be. Typically this is around 85% otherwise the queues become inefficient. \( b \) is a \( 1 \times n \) vector containing the constraint values. In the following section we will describe how this methodology can be applied to the orthopaedic ICATS process to minimise the overall waiting time.

### 3 Orthopaedic ICATS Process

The ICATS process that is described in Section 1 includes all the different elements of the system. However, in this paper we will focus on the five stages where the ICATS team are working. These stages are Paper Triage (1), F2F (2) assessment and the three T&R sessions (3, 4 and 5). In 2009/2010 the number of patients referred to orthopaedic ICATS department was 8522 and this is used to find the external arrival rate to Paper Triage (Southern Health and Social Care Trust, 2010). The internal arrival rates within the department are found using the general traffic equation (Equation 2) (Koizumi, et al. 2005), as follows
\[ \lambda_j = \lambda_i + \sum_{i=1}^{5} r_{ij} \lambda_i \]  

(2)

where \( j = 1, ..., 5 \), \( \lambda_i \) is the arrival rate to stage \( i \), \( \lambda_j \) is the arrival rate to stage \( j \) and \( r_{ij} \) is the proportion of patients who move from stage \( i \) to stage \( j \).

When we have a fixed number of resources, \( C \), which can be distributed throughout the ICATS department we can find the best distribution of resources which minimises the overall waiting time to each stage using the GA described above. Therefore, for our system we want to:

\[
\text{min } f(c_1, ..., c_5) \text{ such that }
\sum_{i=1}^{5} c_i = C, \\
\alpha c_1 \geq \omega_1 \\
\vdots \\
\alpha c_5 \geq \omega_5 \\
c_1, ..., c_5 \geq 0
\]

where \( \alpha \) is the maximum utilization value for each queue which determines how busy the system is allowed to be.

4 Results

We applied the Genetic Algorithm described in Section 2 to two different utilization values of the ICATS department and compared it with the set of solutions that we have found using an Exhaustive Search Algorithm (ESA) programmed in Matlab. An ESA examines every possible solution before selecting the most optimal (Caban and Mersereau, 1998) and the approach is renowned for its long execution times, especially for large datasets (Hui and Yonghui, 2010). Matlab is a numerical computing environment which integrates computation, visualization and programming within a user-friendly interface and presents the solution in a familiar mathematical format. The Maximum Utilization (UMax) values for each stage of the department which we have used are 1 and 0.85 respectively.

Table 1 shows that most of the time the ESA finds a better combination of resources to minimize the overall waiting time. However, the GA finds a very close approximation to the global optimum. We also note that when there are 7 resources and a utilization maximum value of 1 the GA finds a better solution to our problem. We can also see that the ESA takes much longer to run in Matlab compared to the GA as expected.
Table 1. The combinations found using ESA and GA for the ICATS department

<table>
<thead>
<tr>
<th>No. of Resources C</th>
<th>UMax</th>
<th>Method</th>
<th>c₁</th>
<th>c₂</th>
<th>c₃</th>
<th>c₄</th>
<th>c₅</th>
<th>The overall waiting time</th>
<th>Time to perform method (Secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.5</td>
<td>1</td>
<td>ESA</td>
<td>1</td>
<td>2</td>
<td>1.4</td>
<td>1.1</td>
<td>1</td>
<td>1.3932</td>
<td>21.4296</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GA</td>
<td>1.07</td>
<td>1.97</td>
<td>1.28</td>
<td>1.09</td>
<td>1.09</td>
<td>1.6346</td>
<td>2.1469</td>
</tr>
<tr>
<td>0.85</td>
<td>1</td>
<td>ESA</td>
<td>2.2</td>
<td>1.3</td>
<td>1.0</td>
<td>1</td>
<td>1</td>
<td>1.5514</td>
<td>21.2766</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GA</td>
<td>1.01</td>
<td>2.13</td>
<td>1.35</td>
<td>1.01</td>
<td>1</td>
<td>1.4602</td>
<td>2.0954</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>ESA</td>
<td>2.2</td>
<td>1.5</td>
<td>1.2</td>
<td>1.1</td>
<td>1</td>
<td>0.9502</td>
<td>34.0848</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GA</td>
<td>1.01</td>
<td>1.92</td>
<td>1.29</td>
<td>1.23</td>
<td>1.56</td>
<td>1.5426</td>
<td>2.2079</td>
</tr>
<tr>
<td>0.85</td>
<td>1</td>
<td>ESA</td>
<td>2.2</td>
<td>1.5</td>
<td>1.2</td>
<td>1.1</td>
<td>1</td>
<td>0.9502</td>
<td>32.5935</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GA</td>
<td>1.00</td>
<td>2.11</td>
<td>1.27</td>
<td>1.62</td>
<td>1</td>
<td>1.2773</td>
<td>2.1376</td>
</tr>
<tr>
<td>7.5</td>
<td>1</td>
<td>ESA</td>
<td>2.3</td>
<td>1.7</td>
<td>1.3</td>
<td>1.2</td>
<td>1</td>
<td>0.7182</td>
<td>52.9350</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GA</td>
<td>1.04</td>
<td>2.26</td>
<td>1.63</td>
<td>1.33</td>
<td>1.24</td>
<td>0.7125</td>
<td>2.1644</td>
</tr>
<tr>
<td>0.85</td>
<td>1</td>
<td>ESA</td>
<td>2.3</td>
<td>1.7</td>
<td>1.3</td>
<td>1.2</td>
<td>1</td>
<td>0.7182</td>
<td>49.0474</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GA</td>
<td>1.06</td>
<td>2.17</td>
<td>1.60</td>
<td>1.41</td>
<td>1.27</td>
<td>0.7236</td>
<td>2.0809</td>
</tr>
</tbody>
</table>

Orthopaedic ICATS has only 5 stages and at present we only require 6 or 7 resources for the system to reach steady state and produce an acceptable waiting time. However, in a more complex department, such as Accident and Emergency which has to treat a broad spectrum of injuries and diseases, there could be many more stages and hundreds of staff which would leave ESA too costly and time consuming to undertake. Therefore, the GA described above could be used within these circumstances as it is accurate, quicker and does not require nearly as much computational power.

5 Conclusions and Further Work

The orthopaedic Integrated Clinical Assessment and Treatment Service was implemented in the United Kingdom by the Department of Health (DOH) to reduce the bottleneck within the department. However, due to unanticipated arrivals the DOH are aware that queues have started to build up within the new department. In this paper we have used a Genetic Algorithm to distribute a fixed number of resources throughout an orthopaedic ICATS department so that we reduce the queues and minimise the overall waiting time at each stage. We found that although the GA does not always find the global maximum it does find a very close approximation and it less costly and time consuming to run compared with an enumeration method build in Matlab. This methodology is a generic model which could be implemented within any department and would be very beneficial when there are a correspondingly large number of stages and a large number of resources to be distributed.

In further work we would like to implement the GA in different ICATS scenarios to investigate the impact of changing the constraints. This will help to determine how robust our GA is and whether it would be possible to adapt the methodology to implement within a more complex environment. We also plan to
implement the results of our GA within a simulation model to determine the impact of variation and whether we can make the ICATS process more realistic.

References
Application of Parametric Homogeneity of Variances Tests under Violation of Classical Assumption

Alisa A. Gorbunova and Boris Yu. Lemeshko

Novosibirsk State Technical University
Department of Applied Mathematics, Novosibirsk, Russia
(e-mail: gorbunova.alisa@gmail.com, Lemeshko@fpm.ami.nstu.ru)

Abstract. There are a good number of tests that are available for testing a hypothesis that samples come from populations with the same variance. It is well known that classical tests for comparing variances are very sensitive to departures from normality. However, they are more powerful than nonparametric ones. So, the new approach for testing hypotheses of variances homogeneity is proposed. Software for comparing variances using parametric tests (F-test, Cochran’s, Bartlett’s, Hartley’s, Levene’s, modified Levene’s, Neyman-Pearson’s, Z-variance, Overall-Woodward modified Z-variance and O’Brien tests) when samples are from any distribution (skewed, leptokurtic, platykurtic) has been developed. In this case the p-value is defined using a simulated empirical distribution in real-time testing of the hypothesis. Recommendations on choosing the most powerful test for a particular form of data distribution are given.

Keywords: homogeneity of variances, power, simulation study.

Introduction

Testing for equality of variances often attracts attention as preliminary to other analyses involving comparisons of means, such as an analysis of variance (ANOVA) or the t-test. Correct application of tests for means equality implies that variances are equal.

However, preliminary tests of variances equality used before applying a test of location are not recommended by some statisticians. Many authors (e.g., Zar[1]) stated that the tests presently available have such a poor performance that they are not really useful, with ANOVA being more robust to departures from homoscedasticity than can be detected using a test of homogeneity of variances, especially under non-normal conditions. But recent study by Legendre and Borcard[2] has showed that “heterogeneity of variances is always a problem in ANOVA, and is troublesome even in the most benign cases, i.e., when one of the variances is smaller than the others”. So, there is a great need for a test that will correctly detect heterogeneity of variances before applying procedures for means comparison.

The homogeneity of variances tests per se are also of interest in a number of research areas. A variance could be considered as an indicator of uniformity, e.g. in the quality control of manufacturing processes, in agricultural
production systems or in the development of educational methods. Differences in variability of populations could be interesting to biologists, e.g. in the study of genetic diversity or mechanisms of adaption (Boos and Brownie[3]).

So, we want to know whether variances are equal, that is to test hypothesis of variances homogeneity. The null hypothesis for variances equality of m samples has the following form: \( H_0 : \sigma^2_1 = \sigma^2_2 = \ldots = \sigma^2_m \) and the alternative hypothesis is \( H_1 : \sigma^2_i \neq \sigma^2_j \), where the inequality holds at least for one pair of \( i,j \).

To test hypothesis \( H_0 \) there are a good number of available tests both parametric and nonparametric. Moreover, there is considerable statistical literature on testing homogeneity of variances. Therefore, potential user of a test for equality of variances is faced with a confusing array of information concerning which test to use. And worse, this information is sometimes conflicting.

What are the problems when testing equality of variances? Primarily we should choose an appropriate test. Of course, we want to have a robust and powerful test. But it is well known that most parametric (classical) tests for comparing variances are extremely sensitive to the normality assumption. At the same time there are many nonparametric tests that do not depend on sample distribution. But in terms of power all parametric tests have an advantage; they are always significantly more powerful than nonparametric ones.

So, if you want to test the hypothesis of variances equality, you will have to choose between robustness and power. That is why we propose a new approach to testing homogeneity of variances that will help us to avoid problems with the validity of classical tests. In this case the p-value is calculated using a simulated empirical distribution in real-time testing the hypothesis. Then we only should know what test is the most powerful in the particular situation.

Thus, the purpose of this study is to:

- give a possibility to correctly apply parametric tests when the normality assumption may not be true;
- give recommendations on choosing the most powerful test for a particular form of data distribution;
- give recommendations on choosing the best test in terms of robustness in the case of using a software that do not provides simulation.

1 Description of tests studied

A great number of tests for the variances homogeneity have been proposed and examined in statistical literature. But so far, the most frequently cited and used methods have been the \( F \)-test, Bartlett’s, Cochran’s, Hartley’s and Levene’s tests. Below we give a description of these tests.
**F-test.** Let $S_1^2$ and $S_2^2$ denote the variance estimates of samples with sizes $n_1$ and $n_2$ respectively, the classical F-test utilizes the following test statistic:

$$ F = \frac{S_1^2}{S_2^2}. $$

This test statistic has the $F_{n_1-1,n_2-1}$-distribution if the null hypothesis of variances equality is true. The null hypothesis is rejected if the statistic $F$ is either too large or too small.

**Bartlett’s test** was essentially a generalization of the F-test to the several $k > 2$ populations case (Bartlett[4]). The test statistic $B$ involves a comparison of the separate within-group sums-of-squares to the pooled within-group sum-of-squares:

$$ B = (N - k) \ln S_p^2 - \sum_{i=1}^{k} (n_i - 1) \ln S_i^2, $$

where $N = \sum_{i=1}^{k} n_i$,

$$ S_p^2 = \frac{1}{N-k} \sum_{i=1}^{k} (n_i - 1) S_i^2 $$
is the pooled estimate for the variance. The correction factor $C_B$ is calculated as:

$$ C_B = 1 + \frac{1}{3(k-1)} \left( \sum_{i=1}^{k} \frac{1}{n_i-1} - \frac{1}{N-k} \right) $$

and applied to $B$ to obtain a corrected $B_C$ statistic:

$$ B_C = B / C_B. $$

If hypothesis $H_0$ is true and samples are normally distributed, the statistic $B_C$ has approximately the $\chi^2_{k-1}$ distribution.

**Cochran’s test.** The test introduced by Cochran[5] was considerably easier to compute than the tests used at that time:

$$ C = S_{max}^2 / \left( S_1^2 + S_2^2 + \ldots + S_k^2 \right). $$

Unfortunately, the distribution of Cochran’s test statistic depends on the sample size. Tables of critical values for some combinations of the sample sizes $n$ and the number of groups $k$ have been presented by different authors. If the test statistic $C$ is more than the critical value, the null hypothesis $H_0$ is rejected.

Cochran’s test seems to be the best method to detect cases when the variance of one of the groups is much larger than that of the other groups.

**Hartley’s test.** This test was proposed by Hartley[6] in 1950. It is well known as the “$F$-max” test and is very simple to calculate. Its test statistic is just a ratio between the largest and the smallest sample variances:

$$ H = S_{max}^2 / S_{min}^2. $$

It should be noted that Hartley’s test resembles Cochran’s test with a less optimal use of the information available. One can find in the literature tables of critical values created by Hartley. The tables evaluate the test statistic with degrees of freedom $k$ and $n - 1$ (if $n_1 = n_2 = \ldots = n_k = n$). The hypothesis $H_0$ should be rejected if the test statistic $H$ is large.
Levene’s test. In 1960, Levene[7] proposed using the one-way ANOVA $F$ statistic on the variables $Z_{ij} = |X_{ij} - \bar{X}_i|$ as a method for testing equality of variances. The test statistic is given by:

$$L = \left(\frac{(N-k) \sum_{i=1}^{k} n_i (\bar{Z}_i - \bar{Z})^2}{(k-1) \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Z_{ij} - \bar{Z}_i)^2}\right),$$

where $\bar{X}_i$ is the estimated mean of the $i$th sample, $\bar{Z}_i$ is the mean of $Z_{ij}$ for $i$th sample and $\bar{Z}$ is the overall mean of the $Z_{ij}$.

In some descriptions of this test it is said that the statistic $L$ has a $F_{k-1,N-k}$-distribution. Actually, distribution of Levene’s test statistic is not Fisher’s distribution! If sample sizes are less than 40, the distribution of the statistic differs greatly from Fisher’s one. We must take this into account when using this test.

Levene’s test is less sensitive to departures from normality as compared to other classical tests. However, it is less powerful.

Modified Levene’s test. Miller[8] showed that ANOVA on Levene’s variables $|X_{ij} - \bar{X}_i|$ will be asymptotically incorrect if the population means are not equal to the population medians (essentially requiring symmetry). Brown and Forsythe[9] suggested using the sample median instead of the mean in computing $Z_{ij}$ in the Levene’s test statistic. That is $Z_{ij} = |X_{ij} - \tilde{X}_i|$, where $\tilde{X}_i$ is the median of the $i$th sample. This modification allows us to overcome the above problem by centering the variables.

For this study we have also chosen a group of tests that are referred to as the most powerful ones in recent publications. These are Neyman-Pearson’s, Z-variance, Overall-Woodward modified Z-variance and O’Brien tests.

Neyman-Pearson’s test. The test statistic is defined as the ratio of arithmetic mean to the geometric mean of variance estimates:

$$P = \left(\frac{1}{k} \sum_{i=1}^{k} S_i^2\right) / \left(\prod_{i=1}^{k} S_i^2\right)^{\frac{1}{k}}.$$

Th null hypothesis $H_0$ should be rejected, if $P > P_{\alpha,n,k}$, where $P_{\alpha,n,k}$ is a critical value of this test.

Z-variance test. A normal deviation transformation is used to obtain Z-score equivalents of the sample variance. The test statistic proposed by Overall and Woodward[10] is:

$$V = \left(\sum_{i=1}^{k} Z_i^2\right) / (k - 1),$$
where $Z_i = \sqrt{(c_i(n_i-1)S_i^2)/MSE} - \sqrt{c_i(n_i-1)-c_i/2}$, $c_i = 2 + 1/n_i$, $MSE = \left(\sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 \right) / (N-k)$, $N = \sum_{i=1}^{k} n_i$ - total sample size.

If samples are normally distributed and the null hypothesis is true, the statistic $V$ does not depend on the sample size and has approximately the $F_{k-1,\infty}$-distribution.

**Overall-Woodward modified Z-variance test.** As other classical tests the Z-variance test is extremely sensitive to the normality assumption. So, Overall and Woodward[11] conducted a series of studies to determine a $c$ value so that variances of $Z_i$ would remain stable when samples are not normally distributed. Using regression they have found a $c$ value based on the sample size and the kurtosis.

The new $c$ value is calculated by: $c_i = 2.0 \left(\frac{2.9+0.2/n_i}{K}\right)^{1.6(n_i-1.8K+14.7)/n_i}$, where $K$ is the mean of the kurtosis indices for all samples.

The kurtosis index used by Overall and Woodward is:

$$K_i = \left(\sum_{j=1}^{n_i} G_{ij}^4\right) / (n_i - 2), \text{ where } G_{ij} = (X_{ij} - \bar{X}_i) / \sqrt{S_i^2(n_i - 1)/n_i}.$$

Our study has shown that this test remains stable for distributions with different kurtosis indices. However, it is not true for skewness indices.

**O’Brien test.** O’Brien[12] has claimed that his test is a general method that does fairly well for behavioral science data. He also states that this ANOVA-based test is robust to departures from normality.

The O’Brien test statistic is calculated in the following way. First, every raw value $X_{ij}$ is transformed using the formula:

$$V_{ij} = \frac{((n_i - 1.5)n_i(X_{ij} - \bar{X}_i)^2 - 0.5S_i^2(n_i - 1)) / ((n_i - 1)(n_i - 2))}{\sqrt{S_i^2(n_i - 1)/n_i}}.$$

After this transformation the mean of $V$-values will be equal to the variance for original values, that is $\bar{V}_i = \left(\sum_{j=1}^{n_i} V_{ij}\right) / n_i = S_i^2$.

The O’Brien test statistic will be the $F$-value computed applying the usual ANOVA procedure on the transformed values $V_{ij}$. If the null hypothesis of equal variances is true, the test statistic has approximately the $F_{k-1,N-k}$-distribution.

2 Design of simulation study

All tests described above were compared in terms of robustness and power using Monte Carlo studies. The tests studied were (1) $F$-test, (2) Bartlett’s, (3) Cochran’s, (4) Hartley’s ($F$-max), (5) Levene’s, (6) modified Levene’s, (7) Neyman-Pearson’s, (8) Z-variance, (9) Overall-Woodward modified Z-variance and (10) O’Brien tests.
To know how the form of sample distribution influences the performance of tests we have taken five types of distributions of various forms: (1) normal, (2) leptokurtic, (3) platykurtic, (4) moderately skewed and (5) highly skewed.

We have used the exponential family of distributions with the density $De(\theta_0) = f(x; \theta_0, \theta_1, \theta_2) = \theta_0/(2\theta_1 \Gamma(1/\theta_0)) \exp(-(|x - \theta_2|/\theta_1)^{\theta_0})$ to approximate symmetric distributions. The Laplace distribution ($De(1)$) and the distribution with $\theta_0 = 3$ ($De(3)$) were taken as leptokurtic and platykurtic distributions respectively.

The chi-squared distributions with 6 and 3 degrees of freedom ($\chi^2_6$ and $\chi^2_3$) have been chosen for moderately and highly skewed distributions respectively.

To investigate statistics distributions, to calculate percentage points and to estimate the power of tests we used statistical simulation methods and the software developed. Each test statistic was computed 1 000 000 times. Such a value gives a simulation accuracy of 0.001.

To estimate the tests power, we need to simulate statistics distribution when the alternative hypothesis $H_1$ is true. For this purpose we set several competing hypotheses by manipulating the value of the parameter $r = \sigma^2_{\text{max}}/\sigma^2_{\text{min}}$. The larger $r$ is, the more the corresponding populations depart from the hypothesis of equal variances, i.e. the distance between competing hypotheses is larger. A smaller distance makes it more difficult to detect differences in variances. We have considered different distances between competing hypotheses: small ($r = 1.1$), moderate ($r = 1.2$) and large ($r = 1.5$). According to Whedyka and Nelson[13] three basic variance configurations were studied: (1) $k-1$ variances are equal, the last variance is larger, (2) $k-1$ variances are equal, the last variance is smaller, (3) $k-2$ variances are equal, the first variance is smaller and the last variance is larger.

3 Simulation study results

The study of classical tests power has shown that $F$-test, $Z$-variance, Bartlett’s, Cochran’s, Hartley’s, Neyman-Pearson’s and O’Brien tests have equal and the highest power for two normal samples while the power of Levene’s test is much less. This is also true for platykurtic distributions. But for leptokurtic and skewed distributions Levene’s test is more powerful than the other procedures. Furthermore, the modified Levene’s test outperformed the original test in this case.

Bartlett’s, Cochran’s, Hartley’s, Levene’s, Neyman-Pearson’s, O’Brien, $Z$-variance and modified $Z$-variance tests can be applied when the number of samples $k > 2$. In such situations the power of these tests is different. If the normality assumption is true, these tests can be ordered according to the power decrease in the following way:

Cochran’s $\succ$ O’Brien $\succ$ Z-variance $\succ$ Bartlett’s, Neyman-Pearson’s $\succ$ modified Z-variance $\succ$ Hartley’s $\succ$ Levene’s, modified Levene’s.
The tests preference remains the same for platykurtic distributions. When samples are from leptokurtic or skewed distributions, this preference order changes. Now Levene’s test has a greater power with the modified Levene’s test being more powerful than the original one.

It has been mentioned earlier that Cochran’s test is the best method to detect cases when the variance of one of the groups is much larger than that of the other groups (configuration (1) in this study). However, if the variance configuration differs from the aforementioned, the power of Cochran’s test decreases significantly. So, in such situations we should prefer O’Brien, Z-variance, Bartlett’s or Neyman-Pearson’s tests.

Based on the results obtained we have chosen Cochran’s test and have compiled tables of upper percentage points for some non-normal symmetric distributions. These values can be used in situations when distribution from exponential family $De(\theta_0)$ with the appropriate parameter $\theta_0$ is a good model for observable random variables.

The study of tests robustness has shown that Levene’s, modified Levene’s, Overall-Woodward modified Z-variance and O’Brien tests are the best ones.

Let us formulate recommendations on choosing the appropriate test for comparing variances taking into account all results obtained:

- If there is every reason to consider data distribution as symmetric with excess kurtosis equal or less than zero, i.e. mesokurtic or platykurtic, the best choice will be O’Brien or Z-variance tests. Here Cochran’s test could be recommended only for situations when one variance is larger than others;
- If the data distribution is leptokurtic (excess kurtosis is positive, tails are heavy) or skewed, modified Levene’s test should be chosen.

4 Software for testing hypotheses

It is impossible to develop distribution models for all distributions and sample sizes. So, we have developed the software that allows us to correctly apply tests for comparing variances when samples are from any distributions. We can choose any distribution from the list and simulate a distribution of the statistic. Also, we can set a required accuracy of simulation by defining the size of a statistics sample. Then a p-value is calculated using a simulated empirical distribution.

The simulation process is done using parallel computing, so the speed of simulation depends on the number of CPU cores and the required accuracy. It can be claimed that it takes not much time to make a correct decision when testing the hypothesis of equal variances.

Tables of critical values for Cochran’s test and the latest version of the software can be obtained from the authors upon request.
Acknowledgements

This research was supported by the Russian Foundation for Basic Research (project no. 09-01-00056a), by the Federal Agency for Education within the framework of the analytical domestic target program “Development of the Scientific Potential of Higher Schools” and federal target program of the Ministry of Education and Science of the Russian Federation “Scientific and Scientific-Pedagogical Personnel of Innovative Russia”.

References

Integrative Evaluation of Yield and Flow Time under Combined Inspection and Repair Allocation

Shai Goren, Gad Rabinowitz and Yoav Kerner

Ben-Gurion University of the Negev, Beer-Sheva, Israel
Emails: goren@bgu.ac.il, rgadi@exchange.bgu.ac.il, kerneryo@bgu.ac.il

Abstract

We present a model for integrative evaluation of yield and flow time (FT) performance measures, under combined inspection and repair allocation policies, in a stochastic production environment.

We model and solve a typical production stage as a basis for becoming a node of a queueing network in future studies. The stage consists of a single processing machine. The process deteriorates randomly from in-control to out-of-control (with lower yield), and returns back to in-control only by repair. Repair is triggered via inspection of a recently produced item. Certain portions of the plant’s capacities of inspection and repair resources are allocated to the stage, according to the operational policy. Production, inspection and repair are assumed to be non-perfect (i.e. not always successful), and their durations are randomly distributed.

The model is utilized for evaluating the influence of inspection capabilities on the performance measures (yield and FT). We assume a trade-off between Type I and Type II errors and develop a total cost function that can be used for optimizing the inspection facility's working point.

Keywords: Stochastic modeling, inspection capabilities, performance measures

1 Introduction

Modern industries (e.g. semiconductor) are constantly becoming more complex and difficult to control. Customer expectations drive quality requirements to be tighter, and staying profitable requires high operational efficiency. Traditional decision support tools include deterministic (spreadsheet) and simulation models. While the first ones are too naive to capture the manufacturing processes' complexity, the last ones are too time consuming and expensive. Hence, a need arises for a more efficient tool. Analytical models may offer a better trade-off between accuracy and computational efforts.

The performance measures of interest are the yield and the duration items spend in the process, i.e. their FT. These two measures are undoubtedly the most important in most manufacturing plants. A survey of yield models is provided by Kumar et al. (2006), Rabinowitz and Emmons (1997), Emmons and Rabinowitz (2002), and Rabinowitz (2009) studied the yield performance under various inspection and
repair policies. The trade-off between yield and FT was studied by Wein, (1992) Srinivasan et al. (1995) and Tirkel et al. (2009) but under assumptions different than ours.

Recently, in Goren et al. (2012), we developed an analytical stochastic model for the evaluation of yield and FT in a single production stage. In this paper we utilize that model to study the influence of inspection capabilities on these performance measurements.

We chose to focus on the role of inspection capabilities in the stochastic production environment since they are of great interest for practitioners. Setting the control limits of a process, and tuning the inspection facility, are essential stages in the design and control of every manufacturing system, and have a major impact on the system's output.

The capabilities of every inspection facility can be characterized through its inspection errors – Type I (false negative) and Type II (false positive). Naturally, one would prefer both errors to be as small as possible. Yet, for a given inspection effort there is inevitable trade-off between the two types. See Craiu and Sun (2008), Bishop and Winn (2000), Buhl-Mortensen (1996), Harris et al. (2000) and Persson and Siven (2007) for examples in different fields.

In section 2 we briefly describe the single stage model, as a basis for the analysis of inspection capabilities' influence on the performance measures, provided in section 3. Finally, in section 4, we conclude and detail our plans for further development of this model.

2 The Single Stage Model

For the convenience of the readers, we now shortly describe the problem and the model we use to solve it. For more details please refer to Goren et al. (2012).

Figure 1: a typical single production stage

Consider the typical production stage \( j \), illustrated in Figure 1. The stage includes a
single production machine denoted \textit{Process}, with an exponential service rate of \( \mu \) (all rates are in product units, denoted \textit{lots}, per time unit), and an infinite waiting line, denoted \textit{buffer}. A Poisson arrival of \textit{lots} enters the \textit{buffer} at a constant rate of \( \lambda \). The \textit{Process} is either in-control (IC), producing at a yield of \( Y_{IC} \), or out-of-control (OOC), producing at a lower yield of \( Y_{OOC} \). Shifting from IC to OOC occurs randomly, with probability \( p_1 \), at the completion of processing a lot. The process is returned to IC only by a successful repair.

The second component of the stage is the inspection operation, denoted \textit{Inspect}, with an exponential service rate of \( \beta \). \textit{Inspect} tests some of the lots coming out of the \textit{Process}, in order to define the \textit{Process}' state (IC or OOC). We employ the freshest lot (FL) inspection policy suggested by Tirkel et al. (2009). According to the FL policy, the last processed lot is always kept at the freshest lot position (FLP). \textit{Inspect} is subject to both Type I (with probability \( 1-p_2 \)) and Type II (with probability \( p_3 \)) errors. We denote \( P\{T^I\} \) and \( P\{T^II\} \) the probabilities for Type I and Type II errors, respectively.

The third component is the repair operation, denoted \textit{Repair}, with an exponential service rate of \( \varphi \), given the \textit{Process} arrived IC, or \( \psi \), given the \textit{Process} arrived OOC. \textit{Repair} is imperfect; it is successful with probability \( p_4 \), given the \textit{Process} arrived IC, and \( p_5 \), given the \textit{Process} arrived OOC.

Information is assumed imperfect, hence, the true status of the \textit{Process} is never known. Instead, the perceived status, as determined by the inspection results, is used to decide whether the \textit{Process} needs repair.

Central inspection and repair resources are used, and the problem is how to allocate them to the different production stages. We follow the method proposed by Reshef and Rabinowitz (2008) and allocate the inspection and repair resources to stage \( j \) through stochastic exponential rates denoted \( \gamma_j \) and \( \delta_j \) respectively. The \( \gamma \)'s and the \( \delta \)'s are

\[
\begin{align*}
\gamma_1 &= \mu (1-p_1) \\
\delta_1 &= \mu p_1 \\
\gamma_2 &= \mu (1-p_1) \\
\delta_2 &= \mu p_1 \\
\gamma_3 &= \mu (1-p_1) \\
\delta_3 &= \mu p_1 \\
\gamma_4 &= \beta p_2 \\
\delta_4 &= \beta (1-p_2) \\
\gamma_5 &= \beta p_2 \\
\delta_5 &= \beta (1-p_2) \\
\gamma_6 &= \beta p_2 \\
\delta_6 &= \beta (1-p_2) \\
\gamma_7 &= \beta (1-p_3) \\
\delta_7 &= \beta p_3 \\
\gamma_8 &= \beta (1-p_3) \\
\delta_8 &= \beta p_3 \\
\gamma_9 &= \beta (1-p_3) \\
\delta_9 &= \beta p_3 \\
\gamma_{10} &= \psi (1-p_5) \\
\delta_{10} &= \psi p_5 \\
\gamma_{11} &= \psi (1-p_5) \\
\delta_{11} &= \psi p_5 \\
\gamma_{12} &= \psi (1-p_5) \\
\delta_{12} &= \psi p_5
\end{align*}
\]

Figure 2: the single stage state transition mechanism
δ's, i.e. the allocation of the inspection and repair resources to the different production stages, are the decision variables of the multi-stage problem. The single stage state transition mechanism is described in Figure 2. We defined the problem as a Quasi-Birth and Death (QBD) problem. To find its equilibrium distribution we employed Matrix-Geometric methods (Neuts (1994)) and used a standard software package (Mathematica™). The results enable us to calculate the expected values of some important measures such as yield, FT, resources' availability and utilization, and others. In our forthcoming study, multiple stages will be connected to form a serial line and later a network topology. The objective will be to minimize the costs of loss of yield and of FT. Therefore, we are now interested in the yield and the FT of the single stage.

3 Experiment

Design
As explained in the introduction, the focus of this study is on the influence of the inspection capabilities on the performance measures (yield and FT). Inspection capability is determined by the Type I (false negative) and Type II (false positive) errors. We assume that a trade-off exists between the two types of error and we will show how this trade-off affects the performance measures. We follow the above mentioned references and use a non-linear trade-off curve as shown in Figure 3. We chose the values of P{T\text{I}}, i.e. the probability of sending the Process to an unnecessary repair, to be up to 10% because higher values are usually not acceptable. The tradeoff is set by the condition: P{T\text{I}}\cdot P{T\text{II}}=0.1\%. Naturally, in practice, one should characterize the relevant Type I versus Type II trade-off curve, and define the acceptable working range.

![Figure 3: Type I versus Type II trade-off curve](image)

We have tested the influence of the inspection capabilities on the performance measures under various scenarios (see Table 1), in order to see if there is a repeating pattern. The baseline scenario is number 1 and the parameters that differ from it are highlighted in gray. We tested different inspection (γ) and repair (δ)
allocation rates in scenarios 1-4; different process reliability levels ($p_1$) in scenarios 1, 5 and 6; and different service rates ($\mu$) in scenarios 1, 7 and 8.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$p_1$</th>
<th>$1-p_2$</th>
<th>$p_3$</th>
<th>$1-p_4$</th>
<th>$p_5$</th>
<th>$\beta$</th>
<th>$\varphi$</th>
<th>$\psi$</th>
<th>$\lambda$</th>
<th>$\mu$</th>
<th>$\gamma$</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95%</td>
<td>1-10%</td>
<td>95%</td>
<td>90%</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>95%</td>
<td>1-10%</td>
<td>95%</td>
<td>90%</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>95%</td>
<td>1-10%</td>
<td>95%</td>
<td>90%</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>17</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>95%</td>
<td>1-10%</td>
<td>95%</td>
<td>90%</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>90%</td>
<td>1-10%</td>
<td>95%</td>
<td>90%</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>85%</td>
<td>1-10%</td>
<td>95%</td>
<td>90%</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>95%</td>
<td>1-10%</td>
<td>95%</td>
<td>90%</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>13</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 1: tested scenarios

Results
In all of the scenarios we checked the same pattern repeated. The FT is (nearly linearly) increasing as $P\{T^I\}$ increases, and $P\{T^H\}$ decreases accordingly, as shown in Figure 4. This is because the Process is more often sent to unnecessary repair as $P\{T^I\}$ increases, and, simultaneously, for necessary repair as $P\{T^H\}$ decreases (see the state transition mechanism in Figure 2).

![Figure 4: typical FT curve as a function of inspection capabilities](image)

The typical yield curve is shown in Figure 5. In some scenarios the yield increases asymptotically as $P\{T^I\}$ increases.
To better understand the influence of each of the two inspection errors on the yield we ran the model for \( P\{T^I\} \) between 1-10\%, while \( P\{T^{II}\} \) was kept at 5\%, and then for \( P\{T^{II}\} \) between 1-10\%, while \( P\{T^I\} \) was kept at 5\%. The results are provided in Figure 6.

It is now clear that the yield decreases both with \( P\{T^I\} \) and \( P\{T^{II}\} \). Hence, the joint effect of the inspection capabilities on the yield is a result of the trade-off between them, as shown in Figure 5.

In order to assess the combined influence of the inspection capabilities on both the yield and the FT, we shall now refer to the cost of the loss of yield, denoted \( C_{1,Y} \), and to the cost of FT, denoted \( C_{FT} \). The total cost (TC) will be \( TC=C_{1,Y}(1-Y)+C_{FT} \), as shown in Figure 7.

The outcome of this pattern is that TC has a single minimum in the relevant range, representing the trade-off between the costs caused by loss of yield and those caused by the time lots spent in the system (i.e. their FT). The minimal total cost may be at \( P\{T^I\}=1\% \) when \( C_{FT} \) is significantly greater than \( C_{1,Y}(1-Y) \), at \( P\{T^I\}=10\% \) when \( C_{FT} \) is significantly smaller than \( C_{1,Y}(1-Y) \), and anywhere between when \( C_{FT} \) and \( C_{1,Y}(1-Y) \) are of the same magnitude.
4 Conclusions

The single stage model was developed in order to set inspection and repair rates for different stages in a multi-stage system. In this paper we utilize it to choose an optimal working point of the inspection facility along the trade-off curve of Type I and Type II errors. Simultaneous reduction of the two error types, when feasible at all, will require further investment in inspection resources, so that the trade-off curve will move towards the axes’ origin. The extra cost can be added to the objective function, so that it will be part of the optimization.

We have shown how the inspection capabilities affect the performance measures of the single stage. We have tested a few different scenarios and revealed a repeated pattern of the costs’ structure. While the influence of the inspection capabilities can be estimated from analyzing the state transition mechanism, the shape of the costs structure curves cannot. Hence, our analysis provides a non-trivial understanding of the model and its inter-connections.

In our forthcoming studies we will further investigate the influence of other parameters of the single stage model on the performance measures. In addition, we will utilize the single stage as a component of a wider multi-stage model, to make it closer to real production systems.

References

4. Emmons, H. and Rabinowitz, G., Inspection allocation for multistage


Geostatistical reconstruction of sedimentary heterogeneity for groundwater numerical modeling of a contaminated aquifer: the industrial site of Pianvallico (Florence, Italy)

Enrico Guastaldi\textsuperscript{1}, Andrea Carloni\textsuperscript{2}, Enrico Giaquinta\textsuperscript{3}, Fabio N.A. Brogna\textsuperscript{1}, Giovanni Liali\textsuperscript{1}, Carmela Rezza\textsuperscript{1}, and Angela Soggia\textsuperscript{1}

\textsuperscript{1} CGT Center for GeoTechnologies, University of Siena
Via Vetri Vecchi, 34
52027 San Giovanni Valdarno, Arezzo, Italy
(Corresponding Author’s e-mail: guastaldi@unisi.it).
\textsuperscript{2} ENEA
Casaccia, Rome, Italy.
\textsuperscript{3} Terranova S.r.l.
53040, Via Pietro Nenni 9/A, Rapolano Terme, Siena, Italy.

Abstract. Both geostatistical and inverse statistical approach play a growing role in modes hydrogeology. The characterization of sedimentary heterogeneity in subsurface performed by a great number of geological, geophysical, hydrogeological and hydrological information makes more real the predictive capability of a numerical modeling of groundwater flow and pollutant transport. Downstream the environmental characterization study for the reclamation of a contaminated site in the north of Tuscany region (Italy), this work concerns the sedimentary reconstruction of a contaminated aquifer through four stochastic methodologies. These geostatistical simulation methods based on different conceptual theories (structure imitating and process imitating approaches) yielded simulated sedimentary volumes, basis of groundwater flow and transport modeling, which highlight advantages and drawbacks in hydrogeological numerical real application. We applied the results of this study for sizing the remediation project parameters of this contaminated industrial site, demonstrating geostatistical simulations results can be utilized in real case studies only after an inverse modeling procedure that tests what is the more accurate simulator.

Keywords: Geostatistics, Stochastic simulation, Groundwater flow numerical modeling, Truncated Plurigaussian Simulation, Markov Chain, Sequential Indicator Simulation, Turning Bands Simulation.

1 Introduction

One of the main environmental topics is the quantitative analysis of pollutants transport in the groundwater, always related to the geological heterogeneity of porous medium where the contaminated aquifer lies. In fact, the quantification of reservoir and its characteristics not only is an interesting academic research goal, but also it plays a fundamental role in real
environmental problems for designing both the best economic and effective reclamation systems. So, numerical analysis based on both conceptualizing the geological characteristics of a contaminated site and geostatistical simulations that not modify the physical consistence of the analyzed phenomenon, allow to constitutes a valid support for the sustainable economical assessing that considers the variability of initial data.

In particular, the most part of aquifer is made by high hydraulic conductivity sediments that allow groundwater flow in preferential paths network alternating to low hydraulic conductivity bodies, limiting the groundwater flow. In this hydrogeological conditions, solutes transport shows features not ascribable to the advection-dispersion Fick’s law conditions (Carrera et al., 1998 [2]).

---

Fig. 1. A) Study area location; B) Frequency distribution of continuous lengths of four lithologies composing both the upper and lower parts of Pianvallico’s phreatic aquifer (L, Silt; SL, Silty Sand; S, Sand; SG, Sands and Gravels); C) Data locations in terms of types of available geognostic investigations (D, Drilling; DSHP, Dynamic Super Heavy Probing; GT, Ground Testing; VES, Vertical Electrical Sounding. Coordinate System: Gauss-Boaga kilometric national grid); Three-dimensional representation of data (the available geognostic investigations) with the indications of lithologic composition.

This study pays attention to geostatistical simulation methods for reconstructing the geological sedimentary heterogeneity finalized to both groundwater flow and solute transport numerical simulation for a saturated aquifer. For that, we chosen a contaminated test site localized in Florence province (near to Scarperia and San Piero a Sieve villages; Fig. 1.A), widely studied...
in terms of geological, hydrogeological, and environmental characterization. We studied several numerical modeling techniques supported by geognostic investigations (Fig. 1.C) for identifying both the geometry of the reservoir and the spatial variability of hydrodynamic properties at small scale. Moreover, we defined the influence of flow paths network in order to choose the better conceptual approach for solute transport simulating, estimating the extent of pollution plume and performing temporal scenarios of contamination evolution.

2 Materials and methods

The lithostratigraphic model was reconstructed by using a 96 geognostic data set (Fig. 1.C and D), distributed within a test area of 800x600 meters where the contaminated plume was detected from 1991 to 2009 (Fig. 2).

The domain has been discretized with a horizontal grid of 5x5 meters, with 16 layers with a thickness comprised between 0.25 and 1 meter. Geological heterogeneity was classified in 4 separate homogeneous categorical variables (Fig. 1.B). On the base of the observed hydraulic properties within

Fig. 2. Conceptual model of transport and dispersion of solute contamination mass in groundwater, function of time and hydrogeological characteristics of contaminated site, based of the detected contamination from 1991 to 2009 (dashed isolines refer to calculated mean water table).
the study area, vertical and horizontal stratigraphic variations are then determined by using different geostatistical simulation methods to both evaluate different probabilistic solutions and possibly reduce uncertainty concerning the structure of the geological model.

Structure imitating methods (Koltermann and Gorelik, 1996 [7]) and other four different spatial statistic methods have been used: Gaussian simulations as the Turning Bands (TB; Chiles and Delfiner, 1999 [3]) and the Truncated Plurigaussian (TP; Le Loc and Galli, 1997 [8]); and No Gaussian simulations such as the Sequential Indicator Simulation (SSI; Juan et al. [6]) and the Markov Chain (MC; Carle and Fogg, 1996 [1]). Some of these are based on spatial variography analysis (SSI and TB), others on the transition probability (MC; see a step of analysis in Fig. 3), the most complex ones on both of them (TP; see a step of analysis in Fig. 4). Frequency distribution of data and spatial covariance function are the same for all geostatistical solutions and the simulations were parameterized using the same stratigraphic conceptual model (imaging reference).

![Fig. 3. Transition probabilities matrix for upper portion of the phreatic aquifer of Pianvallico contaminated site.](image)

A synthesis model of spatial variability is obtained during post-processing by performing a statistical analysis of 50 simulations for each method. The field of hydraulic conductivity is obtained via geostatistical reconstructions of the aquifer through calibration (process-imitating) of flow and transport numerical models simulating the contaminated plume. We used the MOD-
Fig. 4. A phase of TP analysis for upper portion of the Pianvallico aquifer: A) Vertical Proportion Curves (VPC); B) Lithotypes transition rules and histograms of frequencies for lithotypes; C) Vertical variogram models for VPC; D) Horizontal variogram models for VPC

FLOW code (Harbaugh et al, 2000 [5]) for the flow and MT3DMS (Zheng and Wang, 1998 [9]) code for the transport. In order to determine the relationships between the reconstructed geostatistical simulations we compared results in terms of scale of analysis and mass transfer phenomena, namely, advection-dispersion component (ADMT), and dual-domain models (DDMT).

3 Results and Discussion

The solutions provided by the geostatistical models were analyzed both in terms of using Experimental Semi Variogram (ESV), and in terms of geometric continuity and connections of geological structures. The different heterogeneity configurations show a low linear spatial correlation with lithofacies overlapping of about 30%. TB and TP models have 50% of nugget effect, and SSI and MC models about 80%. The SSI model simulates the lower spatial continuity of the geologic structures with a 60 meters range, and TP model show the maximum value of 180 meters range. The study shows that the geostatistical models based on the spatial covariance as SSI and TB, (covariance based), may be inadequate for reproducing continuous macrostructures, as they generate very erratic patterns due to the difficulty to simulate the ESV for excessive anisotropy between XY and Z spatial discretization. The
MC and TP models, which are based on probability transitions (geologic-based), allow to both better reconstruct and highlight the interconnections of geological macrostructures, but local heterogeneity at the micro-scale are not captured (Fig. 5).

![Image](image-url)

**Fig. 5.** Results of four simulation methods: A) TB; B) SSI; C) MC; D) TP

Geologic-based methods revealed the difficulties in the transposition from the 1D to 3D chains and the uncertainty to chose the ratio between the size of the structures, between the ratios of length to thickness which strongly influence the reconstruction of geological elements at the reduced scale. TB and SSI transport models show a wide and irregular in shape dispersion plume compare to MG or TP, which is more regular and compact. Generally, simulating the plume distribution by DDMT models introduces by a large tailing effect compared to those given by ADMT models and a better matching with measured concentrations.

Once applied to the conductivity fields generated by the MC and TP methods, the ADMT model provides solutions with Channel Flow Path plume and with an overestimation of time needed for mass contamination transfer. The Upscaled DDMT transport model furnishes a partially improved solutions. The transport model solution obtained with the TB conductivity field, which is able to represent both macrostructures and lower scales structures, provides a proper calibrated Preferential Flow Path of plume pattern (Fig. 6). Furthermore, the latter considers both the diffusion/ dispersion
Fig. 6. Simulated particles paths in advective flow through simulated heterogeneity by: A) TB; B) SSI; C) MC; D) TP. Red cells correspond only to simulated SG lithology.
and sorption dynamics with apparently no dependency from the modeling approach selected.

4 Conclusion

We presented a field-size case study of a contaminated aquifer characterized by solutes macro-dispersion transport in a strongly heterogeneous soil. In particular, we focused on on different geostatistical simulation methods to reconstructing the hydraulic conductivity field coupled with two transport modeling approaches. Our results indicate how the choice of geostatistical method for reconstructing the geological structure and hydraulic connection unit network at the different scales is a critical aspect for simulation contaminant transport. Finally, we conclude that the choice of the geostatistical model is the key factor that really determines our ability of explicitly accommodate sorption and diffusion phenomena in transport model and that this is much more important than the choice of the modeling approach itself.

References

Combining Multi-class SVMs with Linear Ensemble Methods that Estimate the Class Posterior Probabilities

Yann Guermeur
LORIA-CNRS
Campus Scientifique, BP 239
54506 Vandœuvre-lès-Nancy Cedex, France
(e-mail: Yann.Guermeur@loria.fr)
November 10, 2011

Running Title: Linear ensemble methods for M-SVMs

Keywords: multi-class support vector machines, ensemble methods, multivariate linear models, class posterior probability estimates, capacity control

Mathematics Subject Classification: 68T10, 62G08
Abstract

Roughly speaking, there is one main model of pattern recognition support vector machine, with several variants of lower popularity. On the contrary, among the different multi-class support vector machines which can be found in literature, none is clearly favoured. On the one hand, they exhibit distinct statistical properties. On the other hand, multiple comparative studies between multi-class support vector machines and decomposition methods have highlighted the fact that each model has its advantages and drawbacks. These observations call for the evaluation of combinations of multi-class support vector machines. In this article, we study the combination of multi-class support vector machines with linear ensemble methods. Their sample complexity is low, which should prevent them from overfitting, and the outputs of two of them are estimates of the class posterior probabilities.

1 Introduction

Most of the models developed for pattern recognition are based on a principle that does not change fundamentally with the number of categories. This is not the case of the support vector machines (SVMs). Initially, Vapnik and his co-authors devised a (1-norm) machine dedicated to the computation of dichotomies [3]. Although bi-class variants of this machine exist that exhibit appealing properties, such as the 2-norm SVM [3] or the least squares SVM [21], their use has remained marginal so far. The first studies dealing with the use of SVMs for multi-category classification report results obtained with decomposition methods [20] involving Vapnik’s machine. Multi-class support vector machines (M-SVMs) were only introduced three years [26].

During the last decade, many M-SVMs and decomposition methods involving bi-class SVMs have been introduced and evaluated (see [6] for a survey). Currently, the attention of the community is focused on four main models of M-SVMs: the model of Weston and Watkins [26], the one of Crammer and Singer [4], the one of Lee, Lin, and Wahba [16] and the M-SVM$^2$ [10]. Although their learning problems all extend in a straightforward way the one of the bi-class SVM, they exhibit distinct properties [22]. In recent years, several comparative studies between M-SVMs and decomposition methods have been published [5, 14]. In short, they establish that in practice, no model is uniformly superior or inferior to the others with respect to the standard criteria: prediction accuracy, sparsity, computational complexity, etc. In accordance with what was predicted by the theory, the behaviours observed are different. Strangely enough, to the best of our knowledge, nobody has tried so far to take benefit of that phenomenon by combining different M-SVMs. Filling this void is the subject of this article. More precisely, we deal with the combination of M-SVMs with two requirements in mind: the outputs must be exploitable class posterior probability estimates and the sample complexity of the combiners must be low. The first requirement is motivated by the will to make the post-processing of the outputs easier. Indeed, it is well known that none of these machines produces outputs from which class posterior
probability estimates can be derived straightforwardly. As for the second requirement, it simply stems from the fact that overfitting is one of the main limiting factors in the field of model combination.

Taking our inspiration from the works of Breiman and Friedman dealing with multivariate regression [2], we propose to combine the post-processed outputs of M-SVMs with linear ensemble methods. These methods are based on a multivariate linear model and differ with respect to their objective function. Their sample complexity can be upper bounded thanks to the use of a $\gamma$-$\Psi$-dimension [7] and two of them generate class posterior probability estimates. Their use requires to post-process beforehand the outputs of the machines so that they are nonnegative and sum to one. This is obtained by applying a polytomous logistic regression. Experimental results are provided to illustrate the gain in prediction accuracy resulting from the combinations.

The organization of the paper is as follows. Section 2 introduces the M-SVMs through our generic model encompassing all of them. Section 3 is devoted to the definition and statistical analysis of the linear ensemble methods. The post-processing applied to the outputs of the M-SVMs prior to their combination is described in Section 4. The experimental results are exposed in Section 5. At last, we draw conclusions and outline our ongoing research in Section 6. To make reading easier, proofs have been gathered in appendix.

## 2 Multi-class support vector machines

The theoretical framework for the M-SVMs is the one of large margin multi-category classifiers [7]. It is summarized below.

### 2.1 Theoretical framework

We consider the case of $Q$-category classification problems with $Q \in \mathbb{N} \setminus \{0, 2\}$. Each object is represented by its description $x \in \mathcal{X}$ and the set $\mathcal{Y}$ of the categories $y$ can be identified with the set of indices of the categories: $[1, Q]$. We assume that $(\mathcal{X}, \mathcal{A})$ and $(\mathcal{Y}, \mathcal{B})$ are measurable spaces and the link between descriptions and categories can be characterized by an unknown probability measure $P$ on the measurable space $(\mathcal{X} \times \mathcal{Y}, \mathcal{A} \otimes \mathcal{B})$. $(X, Y)$ is a random pair with values in $\mathcal{X} \times \mathcal{Y}$, distributed according to $P$. We are given a class $\mathcal{G}$ of functions $g = (g_k)_{1 \leq k \leq Q}$ from $\mathcal{X}$ into $\mathbb{R}^Q$. To each function $g$ in $\mathcal{G}$ corresponds a decision rule $f$ defined as follows:

$$
\forall x \in \mathcal{X}, \begin{cases}
\text{if } \exists k \in [1, Q]: g_k(x) > \max_{l \neq k} g_l(x), \text{ then } f(x) = k \\
\text{else } f(x) = \ast
\end{cases}
$$

where $\ast$ denotes a dummy category introduced to deal with the cases of ex æquo. In that context, the learning problem consists in minimizing over $\mathcal{G}$ the risk $P(f(X) \neq Y)$. In practice, since $P$ is unknown, the risk cannot be used directly as objective function. The optimization process,
called training, makes use of a training sample, i.e., an $m$-sample $D_m = ((X_i, Y_i))_{1 \leq i \leq m}$ made up of independent copies of $(X, Y)$, to infer knowledge on $P$.

2.2 Class of functions and learning problem

Defining a bi-class SVM is simple once the concept of maximum margin hyperplane [23] has been introduced. Unfortunately, this concept does not extend to the multi-class case. This can explain why the first unifying definition for the M-SVMs, introduced with minor differences by several researchers (see for instance [28, 7]), is recent. Its main drawback rests in the fact that it does not cover the class of quadratic loss M-SVMs [10]. In [8], we introduced the first generic model of M-SVM encompassing all the machines of this kind published so far. In the sequel, the M-SVMs are considered as instances of this model. To keep the article self-contained, the rest of the section is devoted to its presentation. All the M-SVMs are kernel machines that operate on a unique class of $\mathbb{R}^Q$-valued functions. The underlying vector space of functions can be endowed with a structure of reproducing kernel Hilbert space (RKHS) of vector-valued functions [25].

**Definition 1 (RKHS of $\mathbb{R}^Q$-valued functions $H_{\kappa,Q}$)** Let $\mathcal{X}$ be a non empty set and $Q \in \mathbb{N} \setminus \{0,2\}$. Let $\kappa$ be a real-valued positive type function on $\mathcal{X}^2$ and let $\tilde{\kappa}$ be the real-valued positive type function on $(\mathcal{X} \times \left[1, Q\right])^2$ deduced from $\kappa$ as follows:

$$\forall (x, x') \in \mathcal{X}^2, \forall (k, l) \in \left[1, Q\right]^2, \quad \tilde{\kappa}((x, k), (x', l)) = \delta_{k,l} \kappa(x, x')$$

where $\delta$ is the Kronecker symbol. For each $(x, k)$ in $\mathcal{X} \times \left[1, Q\right]$, let us define the $\mathbb{R}^Q$-valued function $\tilde{\kappa}^{(Q)}_{x,k}$ on $\mathcal{X}$ by the formula

$$\tilde{\kappa}^{(Q)}_{x,k}(\cdot) = (\tilde{\kappa}((x, k), (\cdot, l)))_{1 \leq l \leq Q}. \quad (1)$$

The RKHS of $\mathbb{R}^Q$-valued functions at the basis of a $Q$-category M-SVM whose kernel is $\kappa$, $\left(H_{\kappa,Q}, \langle \cdot, \cdot \rangle_{H_{\kappa,Q}}\right)$, consists of the linear manifold of all finite linear combinations of functions of the form (1) as $(x, k)$ varies in $\mathcal{X} \times \left[1, Q\right]$, and its closure with respect to the inner product

$$\forall (x, x') \in \mathcal{X}^2, \forall (k, l) \in \left[1, Q\right]^2, \quad \langle \tilde{\kappa}^{(Q)}_{x,k}, \tilde{\kappa}^{(Q)}_{x',l} \rangle_{H_{\kappa,Q}} = \tilde{\kappa}((x, k), (x', l)) .$$

**Proposition 1 (Alternative characterization of $H_{\kappa,Q}$)** Let $\mathcal{X}$ be a non empty set and $Q \in \mathbb{N} \setminus \{0,2\}$. Let $\kappa$ be a real-valued positive type function on $\mathcal{X}^2$ and let $\left(H_{\kappa,Q}, \langle \cdot, \cdot \rangle_{H_{\kappa}}\right)$ be the corresponding RKHS. Then, $H_{\kappa,Q} = H^Q_{\kappa}$. Furthermore, the inner product of $H_{\kappa,Q}$ can be expressed as a function of the inner product of $H_{\kappa}$ as follows:

$$\forall (\bar{h}, \bar{h}') \in H^2_{\kappa,Q}, \quad \bar{h} = (\bar{h}_k)_{1 \leq k \leq Q}, \quad \bar{h}' = (\bar{h}_k')_{1 \leq k \leq Q}, \quad \langle \bar{h}, \bar{h}' \rangle_{H_{\kappa,Q}} = \sum_{k=1}^{Q} \langle \bar{h}_k, \bar{h}_k' \rangle_{H_{\kappa}} .$$
Definition 2 (Class of functions \( \mathcal{H}_{\kappa,Q} \)) Let \( \mathcal{X} \) be a non empty set and \( Q \in \mathbb{N} \setminus \{0,2\} \). Let \( \kappa \) be a real-valued positive type function on \( \mathcal{X}^2 \) and let \( \mathbb{H}_{\kappa,Q} \) be the class of functions derived from \( \kappa \) according to Definition 1. Let \( \{1\} \) be the one-dimensional space of real-valued constant functions on \( \mathcal{X} \). The class of functions at the basis of a \( Q \)-category M-SVM whose kernel is \( \kappa \) is

\[
\mathcal{H}_{\kappa,Q} = \mathbb{H}_{\kappa,Q} \oplus \{1\}^Q = (\mathbb{H}_{\kappa} \oplus \{1\})^Q.
\]

A function \( h \) in \( \mathcal{H}_{\kappa,Q} \) can be written as \( h(\cdot) = \bar{h}(\cdot) + b = (\bar{h}_k(\cdot) + b_k)_{1 \leq k \leq Q} \), where \( \bar{h} = (\bar{h}_k)_{1 \leq k \leq Q} \in \mathbb{H}_{\kappa,Q} \) and \( b = (b_k)_{1 \leq k \leq Q} \in \mathbb{R}^Q \). For \( d_m = ((x_i, y_i))_{1 \leq i \leq m} \in (\mathcal{X} \times [1,Q])^m \), \( \mathbb{R}^{Qm}(d_m) \) denotes the subset of \( \mathbb{R}^{Qm} \) made up of the vectors \( v = (v_i)_{1 \leq i \leq Qm} \) satisfying:

\[
(v_{(i-1)Q+y_i})_{1 \leq i \leq m} = 0_m.
\]

For the sake of simplicity, the components of the vectors of \( \mathbb{R}^{Qm}(d_m) \) are written with two indices, i.e., \( v_{ik} \) in place of \( v_{(i-1)Q+k} \) for \( (i,k) \in \[1,m\] \times \{1, Q\} \). Thus, (2) simplifies into \( (v_{ik})_{1 \leq i \leq m} = 0_m \). For \( n \in \mathbb{N}^* \), let \( \mathcal{M}_{n,m}(\mathbb{R}) \) be the algebra of \( n \times n \) matrices over \( \mathbb{R} \) and \( \mathcal{M}_{Qm,Qm}(d_m) \) the subset of \( \mathcal{M}_{Qm,Qm}(\mathbb{R}) \) made up of the matrices \( M = (m_{tu})_{1 \leq t,u \leq Qm} \) satisfying:

\[
\forall j \in \[1,m\], \quad (m_{t,(j-1)Q+y_j})_{1 \leq t \leq Qm} = 0_{Qm}.
\]

Our generic model of M-SVM is defined as follows.

Definition 3 (Generic model of M-SVM, Definition 4 in [8]) Let \( \mathcal{X} \) be a non empty set and \( Q \in \mathbb{N} \setminus \{0,2\} \). Let \( \kappa \) be a real-valued positive type function on \( \mathcal{X}^2 \). Let \( \mathbb{H}_{\kappa,Q} \) and \( \mathcal{H}_{\kappa,Q} \) be the two classes of functions induced by \( \kappa \) according to Definitions 1 and 2. Let \( P_{\mathbb{H}_{\kappa,Q}} \) be the orthogonal projection operator from \( \mathcal{H}_{\kappa,Q} \) onto \( \mathbb{H}_{\kappa,Q} \). For \( m \in \mathbb{N}^* \), let \( d_m = ((x_i, y_i))_{1 \leq i \leq m} \in (\mathcal{X} \times [1,Q])^m \) and \( \xi \in \mathbb{R}^{Qm}(d_m) \). A \( Q \)-category M-SVM with kernel \( \kappa \) and training set \( d_m \) is a large margin discriminant model trained by solving a convex quadratic programming problem of the form

Problem 1 (Learning problem of an M-SVM, primal formulation)

\[
\min_{h, \xi} \left\{ \|M\xi\|^2_p + \lambda \|P_{\mathbb{H}_{\kappa,Q}}h\|^2_{\mathbb{H}_{\kappa,Q}} \right\}
\]

\[
\text{s.t.} \quad \begin{align*}
&\forall i \in \{1,m\}, \forall k \in \{1, Q\} \setminus \{y_i\}, \quad K_1 h_{y_i}(x_i) - h_k(x_i) \geq K_2 - \xi_{ik} \\
&\forall i \in \{1,m\}, \forall (k,l) \in ((\mathcal{X} = [1, Q]\setminus \{y_i\})^2, \quad K_3 (\xi_{ik} - \xi_{il}) = 0 \\
&\forall i \in \{1,m\}, \forall k \in \{1, Q\} \setminus \{y_i\}, \quad (2 - p)\xi_{ik} \geq 0 \\
&1 - K_1 \sum_{k=1}^Q h_k = 0
\end{align*}
\]

where \( \lambda \in \mathbb{R}^*_+ \), \( M \in \mathcal{M}_{Qm,Qm}(d_m) \) is a matrix of rank \((Q - 1)\) \( m, p \in \{1,2\} \), \( (K_1, K_3) \in \{0,1\}^2 \), and \( K_2 \in \mathbb{R}^*_+ \). If \( p = 1 \), then \( M \) is a diagonal matrix.
2.3 Discussion

The reformulations of the learning problems of the four main M-SVMs evoked in introduction as instances of Problem 1 can be found in [8]. Looking at this problem, it appears clearly that even though the concept of maximum margin hyperplane is no longer at the origin of the M-SVMs, the meaning of the values taken by their outputs remains geometrical. Indeed, denoting \( H_{k,l} \) the hyperplane separating categories \( k \) and \( l \) in the RKHS \( H_\kappa \), for all \( x \) in \( X \), the distance between \( \kappa_x = \kappa(x, \cdot) \) and \( H_{k,l} \) is given by:

\[
d(\kappa_x, H_{k,l}) = \frac{|h_k(x) - h_l(x)|}{\|h_k - h_l\|_{H_\kappa}}.
\]

In contrast, the values of the outputs do not provide directly class posterior probability estimates.

3 Linear ensemble methods

For all \( n \) in \( \mathbb{N}^* \), let \( U_n \) be the unit \((n-1)\)-simplex, i.e., \( U_n = \{u = (u_p)_{1 \leq p \leq n} \in \mathbb{R}_+^n : \sum_{p=1}^n u_p = 1\} \).

In this section, we make the hypothesis that \( N \) classifiers (functions from \( X \) into \( \mathbb{R}^Q \)) are available to perform the classification task of interest. Their outputs are supposed to be nonnegative and sum to one, so that they belong to \( U_Q \). We first give a general description of the multivariate linear model (MLM) at the basis of the combiners considered. The linear ensemble methods (LEMs) are obtained by minimizing over this class of functions different objective functions derived from convex loss functions.

3.1 Multivariate linear model

Let \( g^{(j)} = (g^{(j)}_k)_{1 \leq k \leq Q} \) be the \( j^{th} \) classifier. Let \( \tilde{g} \) denote the function from \( X \) into \( U_Q^N \) obtained by appending the component functions of the classifiers \( g^{(j)} \), i.e., \( \tilde{g} = (g^{(j)})_{1 \leq j \leq N} \). Precisely, \( g^{(j)}_k \) is its component function of index \((j-1)Q + k\).

**Definition 4 (Multivariate linear model)** We consider the multivariate linear model parameterized by the matrix \( B \in \mathcal{M}_{Q,NQ}(\mathbb{R}) \) such that

\[
\forall x \in X, \ g_B(x) = B\tilde{g}(x)
\]

s.t. \( \forall v \in U_Q^N, \ Bv \in U_Q \).

This model generalizes the convex combination \( g_\Theta = \sum_{j=1}^N \Theta_{c,j} g^{(j)} \) with \( \Theta_c = (\Theta_{c,j})_{1 \leq j \leq N} \in U_N \). The transposes of the rows of \( B \) are denoted \( \beta_k \), so that the model can be rewritten as:

\[
\forall x \in X, \ \forall k \in {1, \ldots, Q}, \ (g_B)_k(x) = g_{\beta_k}(x) = \beta_k^T \tilde{g}(x). \tag{3}
\]
We denote \( \beta = (\beta_k)_{1 \leq k \leq Q} \in \mathbb{R}^{NQ^2} \) and use alternatively \( g_\beta \) to designate \( g_B \). For the sake of interpretability, the general term of \( B \) (or \( \beta \)) is written with three indices, i.e., \( \beta_{kjl} \) (\( \beta_{kjl} \) is the component of vector \( \beta_k \) of index \( (j-1)Q + l \)). This provides a simple reformulation of (3):

\[
\forall x \in \mathcal{X}, \forall k \in [1, Q], \ g_{\beta_k}(x) = \sum_{j=1}^{N} \sum_{l=1}^{Q} \beta_{kjl} g_{(j)}(x).
\]

The convex combination is the degenerate case obtained by setting \( \beta_{kjl} = \Theta_{c,j} \delta_{k,l} \). We define:

\[
\forall (k,j) \in [1, Q] \times [1, N], \ \beta'_{kj} = \min_{1 \leq l \leq Q} \beta_{kjl}.
\]

For all \( n \in \mathbb{N}^* \), let \( \mathbf{1}_n \) be the vector of \( \mathbb{R}^n \) whose components are all equal to 1. The constraint \( \forall v \in U_N^Q, \ Bv \in U_Q \) defines a convex polytope in \( \mathbb{R}^{NQ^2} \).

**Proposition 2** The expression of the system of constraints of the MLM as a function of the components of matrix \( B \) (vector \( \beta \)) is:

\[
\begin{align*}
\forall k \in [1, Q], & \quad \sum_{j=1}^{N} \beta'_{kj} \geq 0 \\
\forall j \in [1, N], & \quad \forall l \in [1, Q-1], \quad \sum_{k=1}^{Q} (\beta_{kjl} - \beta_{kjQ}) = 0 \\
1^T_{NQ^2} \beta & = Q
\end{align*}
\]

### 3.2 Generic definition of a linear ensemble method

We consider LEMs corresponding to choosing the matrix \( B \) as a sample-based minimizer of a convex risk functional subject to the constraints of Proposition 2.

**Definition 5 (Linear ensemble method)** Given a convex loss function \( \ell_{\text{LEM}} \), a linear ensemble method trained on \( d_m \) is an instance of the MLM whose matrix of parameters, \( B^* \), is obtained by solving the following convex programming problem:

**Problem 2**

\[
\min_B \sum_{i=1}^{m} \ell_{\text{LEM}}(\tilde{g}(x_i), y_i, B) \quad \text{s.t.} \quad \forall v \in U_N^Q, \ Bv \in U_Q.
\]

**Remark 1** Problem 2 is underdetermined. This is due to the fact that the predictors \( g_{(j)}(x) \) are linearly dependent. Let \( \left( U_N^Q \right)^\perp = \left\{ w \in \mathbb{R}^{NQ} : \forall v \in U_N^Q, \ w^T v = 0 \right\} \). If \( \beta^* \) is an optimal solution of Problem 2 and \( \gamma \in \left\{ \left( U_N^Q \right)^\perp \right\} \), then \( \beta^* + \gamma \) is also an optimal solution of Problem 2.

Since we are only looking for one optimal solution of Problem 2, the linear dependency of the predictors can be turned into an advantage by making use of Remark 1 to apply a restriction on the feasible region keeping the quality of the optima unchanged while simplifying computation. The restriction systematically considered in the sequel is characterized by Proposition 3.
Proposition 3 Irrespective of the nature of $\ell_{\text{LEM}}$, there is an optimal solution of Problem 2 which belongs to the nonnegative hyperorthant, i.e., to the convex polytope $V_{N,Q}$ given by:

$$
\begin{cases}
\beta \in \mathbb{R}^{NQ^2}_+ \\
\forall j \in [1, N], \ \forall l \in [1, Q - 1], \ \sum_{k=1}^Q (\beta_{kjl} - \beta_{kjQ}) = 0 \\
1^T_{NQ^2} \beta = Q
\end{cases}
$$

An additional benefit of this restriction is that the set of constraints is directly expressed in standard form. Above all, it makes it possible to characterize the LEMs as implementing a two-level weighting of the predictors. This weighting is defined by Proposition 4.

Proposition 4 (Alternative characterization of $V_{N,Q}$) A vector $\beta$ in $\mathbb{R}^{NQ^2}$ belongs to $V_{N,Q}$ if and only if there exists a vector $\Theta = (\Theta_j)_{1 \leq j \leq N}$ in $U_N$ and a vector $\theta$ in $[0, 1]^{NQ^2}$ satisfying

$$
\forall (j, l) \in [1, N] \times [1, Q], \ (\theta_{kjl})_{1 \leq k \leq Q} \in U_Q
$$

such that

$$
\forall (j, l) \in [1, N] \times [1, Q], \ (\beta_{kjl})_{1 \leq k \leq Q} = \Theta_j (\theta_{kjl})_{1 \leq k \leq Q}.
$$

This proposition highlights the difference between an LEM (with $\beta$ in $V_{N,Q}$) and a convex combination. With a convex combination, each predictor $g_j^{(i)}(x)$ gives its “vote” to the corresponding category $(\theta_{kjl} = \delta_{k,l})$, and this vote is weighted by the weight of the corresponding classifier, $\Theta_{c,j}$. With an LEM, each predictor can split its vote between the different categories $((\theta_{kjl})_{1 \leq k \leq Q} \in U_Q)$. The introduction of this degree of freedom can affect the weighting of the classifiers, i.e., for a given loss function, one can have $\Theta^* \neq \Theta_c^*$.

### 3.3 Estimation of the class posterior probabilities

In this subsection, we focus on two natural choices for the loss function $\ell_{\text{LEM}}$ that give rise to class posterior probability estimates. For all $k$ in $[1, Q]$, let $t_k$ denote the one of $Q$ coding of category $k$, i.e., $t_k = (\delta_{k,i})_{1 \leq i \leq Q}$. The quadratic loss $\ell_{\text{Quad}}$ is defined as:

$$
\forall (x, y, \beta) \in X \times Y \times V_{N,Q}, \quad \ell_{\text{Quad}}(\tilde{g}(x), y, \beta) = \frac{1}{2} \| t_y - g_{\beta}(x) \|_2^2.
$$

Let $\hat{G}$ be the matrix of $\mathcal{M}_{m, NQ}(\mathbb{R})$ such that its row of index $i$ is the transpose of the vector $\tilde{g}(x_i)$. For all $k$ in $[1, Q]$, let $y_k = (\delta_{y,k})_{1 \leq i \leq m} \in \{0, 1\}^m$ and let $y = (y_k)_{1 \leq k \leq Q} \in \{0, 1\}^{Qm}$. The objective function corresponding to the quadratic loss is given by:

$$
J_{\text{Quad}}(\beta) = \frac{1}{2} \beta^T \left\{ I_Q \otimes \left( \hat{G}^T \hat{G} \right) \right\} \beta - \left\{ y^T \left( I_Q \otimes \hat{G} \right) \right\} \beta
$$

where $\otimes$ denotes the Kronecker product. The minimization of $J_{\text{Quad}}(\beta)$ subject to $\beta \in V_{N,Q}$ is a convex quadratic programming problem. The expression of the cross-entropy loss $\ell_{\text{CE}}$ is:

$$
\forall (x, y, \beta) \in X \times Y \times V_{N,Q}, \quad \ell_{\text{CE}}(\tilde{g}(x), y, \beta) = -\sum_{k=1}^Q \delta_{y,k} \ln (g_{\beta_k}(x)).
$$
The expression of the corresponding objective function is

\[ J_{CE}(\beta) = -\sum_{i=1}^{m} \sum_{k=1}^{Q} \delta_{y_i, k} \ln \left( \beta_k^T \hat{g}(x_i) \right). \]

The minimization of \( J_{CE}(\beta) \) subject to \( \beta \in V_{N,Q} \) is a convex programming problem.

To specify the way those two LEMs estimate the class posterior probabilities, we need to introduce additional notations. For all \( x \) in \( \mathcal{X} \), let \( P_x \) be the probability measure on \( \mathcal{Y} \) given by:

\[ \forall k \in [1, Q], \ P_x(k) = P(k|x). \]

For all \( x \) in \( \mathcal{X} \) and all \( \beta \) in \( V_{N,Q} \), let \( P_{x,\beta} \) be the probability measure on \( \mathcal{Y} \) given by:

\[ \forall k \in [1, Q], \ P_{x,\beta}(k) = g_{\beta_k}(x). \]

Furthermore, let \( D_{KL} \) denote the Kullback-Leibler divergence. The aforementioned specification is provided by the following proposition.

**Proposition 5** Irrespective of \( \ell_{LEM} \), let \( \beta^* (m) = (\beta_{k}^*(m))_{1 \leq k \leq Q} \) in \( V_{N,Q} \) be an optimal solution of Problem 2 with \( D_m \) as training sample. Then if the loss function is the quadratic one,

\[
\mathbb{E}_X \left\{ \sum_{k=1}^{Q} \left[ P(k|X) - g_{\beta_k^*(m)}(X) \right]^2 \right\} \xrightarrow{P \to +\infty} \inf_{\beta \in V_{N,Q}} \mathbb{E}_X \left\{ \sum_{k=1}^{Q} \left[ P(k|X) - g_{\beta_k}(X) \right]^2 \right\}, \tag{4}
\]

whereas with the cross-entropy loss we get

\[
\mathbb{E}_X \left[ D_{KL}(P_X \| P_{X,\beta^*(m)}) \right] \xrightarrow{P \to +\infty} \inf_{\beta \in V_{N,Q}} \mathbb{E}_X \left[ D_{KL}(P_X \| P_{X,\beta}) \right]. \tag{5}
\]

The proof of Proposition 5 is inspired from the proofs of similar results obtained for neural networks (see for instance [19]). There are however two fundamental differences regarding the asymptotic behaviour. On the one hand, the classes of functions at the basis of the neural networks considered are universal approximators [12], unlike the MLM. This means that in the first case, the approximation error could be null, whereas in the second case, it should be positive.

On the other hand, an advantage of the LEMs over the aforementioned neural networks, for which the training algorithm may get stuck in local (suboptimal) minima, is that since Problem 2 is a convex programming problem, the training procedure systematically produces an optimal solution. In other words, the estimation error should asymptotically be null. This analysis calls for a justification of the choice of an LEM. As was pointed out in the introduction, it rests on our concern to devise combiners of low capacity, with the aim to avoid overfitting. We conjecture that in many practical cases of pattern recognition, the capacity of a combiner should be superior to that of a simple convex combination and inferior to that of a neural network (or an M-SVM).

The quality of the class posterior probability estimates is primarily governed by three factors: the number of classifiers combined, the nature of their outputs, and the correlation of their errors. Obviously, this quality will be all the better as the predictors \( g^{(j)}_k(x) \) are themselves good estimates of these probabilities. This should be taken into account when processing the outputs of the M-SVMs so as to obtain vectors in \( U_Q \) (see Section 4).
3.4 Model selection with the $\ell_1$ norm

The main advantage of the choice of the $\ell_1$ norm in place of the $\ell_2$ one is well known in machine learning: it leads to sparse solutions. In the framework of our study, instantiating $\ell_{\text{LEM}}$ with

$$\forall (x, y, \beta) \in \mathcal{X} \times \mathcal{Y} \times V_{N,Q}, \quad \ell_{\ell_1} (\tilde{g}(x), y, \beta) = \| t_y - g_{\beta}(x) \|_1$$

produces the following linear programming problem:

**Problem 3**

$$\min_{\beta} \left\{ - \left[ \sum_{i=1}^{m} (t_{yi} \otimes \tilde{g}(x_i)) \right]^T \beta \right\}$$

s.t. $\beta \in V_{N,Q}$.

The sparsity of the optimal solutions of Problem 3 can be characterized exactly. This calls for the characterization of the extreme points of $V_{N,Q}$.

**Lemma 1 (Extreme point of $V_{N,Q}$)** $\beta \in V_{N,Q}$ is an extreme point of $V_{N,Q}$ if and only if

$$\exists j_0 \in \{ 1, N \}, \begin{cases} \forall l \in [1, Q], \exists k_0 (l) \in [1, Q]: (\beta_{kj0(l)})_{1 \leq k \leq Q} = t_{k0(l)} \end{cases}$$

(6)

**Remark 2** A vector $\beta_0$ in $V_{N,Q}$ is an extreme point of $V_{N,Q}$ if and only if

$$\| \beta_0 \|_2 = \max_{\beta \in V_{N,Q}} \| \beta \|_2 = \sqrt{Q}.$$ 

The following proposition is a direct consequence of Lemma 1.

**Proposition 6** There exists an optimal solution $\beta^* = (\beta_{k}^*)_{1 \leq k \leq Q}$ of Problem 3 that can be characterized as follows: there exists $j_0$ in $[1, N]$ and a map $k_0$ from $[1, Q]$ to itself such that

$$\forall k \in [1, Q], \quad g_{\beta^*_k} = \sum_{l:k_0(l)=k} g^{(j_0)}_l.$$ 

(7)

**Corollary 1** Let us consider the LEM whose learning problem is specified by Problem 3. Keeping the notations of Proposition 6, it appears that except in pathological cases, the map $k_0$ should be the identity, with the consequence that $g_{\beta^*_k} = g^{(j_0)}$: the LEM is in fact a model selection method.

3.5 Sample complexity of the linear ensemble methods

To compute the sample complexity of the LEMs, we derive an upper bound on the capacity of the MLM. To that end, we consider an extended definition of this model, corresponding to changing its domain for $U_{N,Q}^N$. Indeed, the bound only depends on the predictor vector through the
domain in which it takes its value (irrespective of the nature of its components). The extension of Definition 4 is thus:

\[ \forall \beta = (\beta_k)_{1 \leq k \leq Q} \in V_{N,Q}, \quad g_\beta : U_Q^N \rightarrow U_Q \quad v \mapsto g_\beta(v) = (\beta_k^T v)_{1 \leq k \leq Q}. \]

In [7], we enriched the Vapnik-Chervonenkis (VC) theory of large margin multi-category classifiers by proving that for the classes of functions at the basis of these classifiers, the appropriate generalizations of the standard capacity measure of the binary models, the VC dimension [24], are the \( \gamma \)-\( \Psi \)-dimensions. Their use is based on the application of margin operators. The operator needed to characterize the capacity of the MLM is the \( \Delta \) one.

**Definition 6 (\( \Delta \) operator)** Let \( \mathcal{G} \) be a class of functions on a set \( \mathcal{X} \) taking their values in \( \mathbb{R}^Q \). \( \Delta \) is defined as an operator on \( \mathcal{G} \) such that:

\[ \Delta : \mathcal{G} \rightarrow \Delta \mathcal{G}, \quad g \mapsto \Delta g = ((\Delta g)_k)_{1 \leq k \leq Q} \]

\[ \forall x \in \mathcal{X}, \quad \Delta g(x) = \frac{1}{2} \left( g_k(x) - \max_{i \neq k} g_i(x) \right)_{1 \leq k \leq Q}. \]

The \( \gamma \)-\( \Psi \)-dimension used is a scale-sensitive extension of the Natarajan dimension [17].

**Definition 7 (Natarajan dimension with margin \( \gamma \))** Let \( \mathcal{G} \) be a class of functions on a set \( \mathcal{X} \) taking their values in \( \mathbb{R}^Q \). For \( \gamma \) in \( \mathbb{R}_+^*, \) a subset \( s_{\mathcal{X}^n} = \{ x_i : 1 \leq i \leq n \} \) of \( \mathcal{X} \) is said to be \( \gamma \)-N-shattered by \( \Delta \mathcal{G} \) if there is a set \( I(s_{\mathcal{X}^n}) = \{ (i_1(x_i), i_2(x_i)) : 1 \leq i \leq n \} \) of \( n \) pairs of distinct indices in \([1,Q]\) and a vector \( \mathbf{c} = (c_i)_{1 \leq i \leq n} \) in \( \mathbb{R}^n \) such that, for each vector \( \mathbf{y} = (y_i)_{1 \leq i \leq n} \) in \([-1,1]^n\), there is a function \( g_\mathbf{y} \) in \( \mathcal{G} \) satisfying

\[ \forall i \in [1,n], \quad \begin{cases} \text{if } y_i = 1, & (\Delta g_\mathbf{y})_{i_1(x_i)}(x_i) - c_i \geq \gamma \\ \text{if } y_i = -1, & (\Delta g_\mathbf{y})_{i_2(x_i)}(x_i) + c_i \geq \gamma \end{cases}. \]

The Natarajan dimension with margin \( \gamma \) of the class \( \Delta \mathcal{G} \), \( N\text{-dim}(\Delta \mathcal{G}, \gamma) \), is the maximal cardinality of a subset of \( \mathcal{X} \) \( \gamma \)-N-shattered by \( \Delta \mathcal{G} \), if this cardinality is finite. If no such maximum exists, \( \Delta \mathcal{G} \) is said to have infinite Natarajan dimension with margin \( \gamma \).

Let \( \mathcal{G}_\beta = \{ g_\beta : \beta \in V_{N,Q} \} \). An upper bound on \( N\text{-dim}(\Delta \mathcal{G}_\beta, \gamma) \) is provided by Theorem 1.

**Theorem 1 (Upper bound on the capacity of the MLM)**

\[ \forall \gamma \in \left( 0, \frac{1}{2} \right], \quad N\text{-dim}(\Delta \mathcal{G}_\beta, \gamma) \leq \left( \frac{Q}{2} \right)^{\frac{NQ}{4\gamma^2}}. \quad (8) \]

For \( \gamma > \frac{1}{4} \), \( N\text{-dim}(\Delta \mathcal{G}_\beta, \gamma) = 0 \). Theorem 1 thus deals with the nontrivial case. In conjunction with Theorem 4.1 in [7], it can be used to optimize the split of the training set into data used to train the M-SVMs, their post-processing, and the selected LEM. An alternative solution consists in considering instead the Rademacher complexity of the different models. In order to derive the corresponding guaranteed risks, one can adapt the proof of Theorem 5.3 in [7].
Figure 1: Flowchart of the computation of the outputs of an LEM. For each level, the set of possible training criteria is mentioned on the left.

4 Implementation of the combinations

Since the M-SVMs do not take their values in $U_Q$, their outputs must be post-processed prior to being combined. This post-processing is all the more important as it performs a transition between models based on different notions of risk. The objective function of an M-SVM is more directly related to the ultimate criterion, the recognition rate, than the objective function of an LEM (using the quadratic loss or the cross-entropy loss) with the problematic consequence that there is no guarantee that the combination should improve this rate. This calls for the choice of a post-processing maximizing the correlation between the behaviours assessed by means of the different measures of prediction accuracy involved. The solution we propose is a variant of the polynomalous logistic regression model [13]. Thus, the functional dependency between an input $x \in \mathcal{X}$ and the corresponding output of an LEM, $g_{\beta}(x) \in U_Q$, is represented by Figure 1.

The predictors of the LEMs take the form

$$ g^{(j)}_{k}(x) = \exp\left( a_{jk}h^{(j)}_{k}(x) + b_{jk} \right) \sum_{l=1}^{Q} \exp\left( a_{jl}h^{(j)}_{l}(x) + b_{jl} \right). $$

This formulation is used to emphasize the fact that the classifiers $g^{(j)}$ actually take their values in $U_Q$. A maximum likelihood estimation of the $N$ models is performed. The outputs $g^{(j)}_{k}(x)$ are initial estimates of the class posterior probabilities, which is a useful feature given the specifications of the LEMs. It is noteworthy that this model can also be seen as a multivariate extension of Platt’s model [18] consisting in fitting a sigmoid after a bi-class SVM.
5 Experimental results

Protein secondary structure prediction is an open problem of central importance in predictive structural biology. It consists in assigning to each residue (amino acid) of a protein sequence its conformational state. We consider here a three-state description of this structure, with the categories being: $\alpha$-helix, $\beta$-strand and coil.

To assess the combinations of M-SVMs with LEMs on this problem, we used a well-known benchmark: CB513. The 513 sequences of this set are made up of 84119 residues. Each sequence is represented by a position-specific scoring matrix (PSSM) produced by PSI-BLAST. The predictor vector used by the M-SVMs corresponds to the content of a window sliding on the PSSMs. Let $2t + 1$ be the size of this window. The predictor vector for the residue of index $n$ in a given sequence is obtained by appending the rows of the corresponding PSSM whose indices range from $n - t$ to $n + t$. The LEMs also use a sliding window on the outputs of the M-SVMs. In practice, introducing this second window is equivalent to increasing the number of base classifiers from $N$ to $N$ times the size of the window. In this configuration, the architecture depicted in Figure 1 applies the standard cascade approach. Precisely, the M-SVMs and the logistic regression models perform the sequence-to-structure prediction whereas the LEM performs the structure-to-structure prediction (filtering).

The M-SVMs combined are the four main models evoked in introduction. They use a dedicated Gaussian kernel derived from the one introduced in [9]. It applies a weighting on the positions of the sliding window, so as to optimize the influence of the local context on the prediction. Both windows are centered on the residue of interest and their respective sizes are 15 (sequence-to-structure) and 17 (structure-to-structure). To assess performance, a ten-fold cross-validation procedure was applied. At each step, the training set is divided as follows: 6/9 for the M-SVMs, 1/9 for the logistic regression models, and 2/9 for the LEM. To compare the prediction accuracy of the LEMs, the combination was also performed with the polytomous logistic regression model and an MLP. In that case, since no post-processing of the outputs of the M-SVMs was required, the combiners were trained one third ($1/9 + 2/9$) of the training set (instead of 2/9). At last, the baseline performance was provided by a cascade of two MLPs.

The secondary structure prediction is often used to provide constraints for tertiary structure prediction methods. This implies that a secondary structure prediction must fulfill different requirements to be useful for the biologist. Thus, several standard measures giving complementary indications must be used to assess the prediction accuracy [1]. We selected the three most popular ones: the recognition rate $Q_3$, Pearson-Matthews correlation coefficients $C_{\alpha/\beta/coil}$, and the segment overlap measure ($\text{Sov'}99$). The quality of the probability estimates was measured by means of the (averaged) cross-entropy (CE). Table 1 summarizes the results obtained.

Using the two sample proportion test, the comparison of the performance of the M-SVMs
<table>
<thead>
<tr>
<th>Architecture</th>
<th>$Q_3$ (%)</th>
<th>$C_{\alpha}$</th>
<th>$C_{\beta}$</th>
<th>$C_{\text{coil}}$</th>
<th>Sov'99</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>WW-M-SVM</td>
<td>77.4</td>
<td>0.73</td>
<td>0.62</td>
<td>0.60</td>
<td>71.8</td>
<td>-</td>
</tr>
<tr>
<td>CS-M-SVM</td>
<td>77.5</td>
<td>0.72</td>
<td>0.63</td>
<td>0.60</td>
<td>72.2</td>
<td>-</td>
</tr>
<tr>
<td>LLW-M-SVM</td>
<td>76.8</td>
<td>0.71</td>
<td>0.62</td>
<td>0.59</td>
<td>71.2</td>
<td>-</td>
</tr>
<tr>
<td>M-SVM$^2$</td>
<td>77.4</td>
<td>0.72</td>
<td>0.63</td>
<td>0.59</td>
<td>71.9</td>
<td>-</td>
</tr>
<tr>
<td>cascade of MLPs</td>
<td>75.4</td>
<td>0.70</td>
<td>0.60</td>
<td>0.56</td>
<td>72.1</td>
<td>0.585</td>
</tr>
<tr>
<td>M-SVMs + Log. reg.</td>
<td>78.0</td>
<td>0.73</td>
<td>0.63</td>
<td>0.59</td>
<td>74.7</td>
<td>0.550</td>
</tr>
<tr>
<td>M-SVMs + MLP</td>
<td>78.2</td>
<td>0.74</td>
<td>0.64</td>
<td>0.59</td>
<td>73.8</td>
<td>0.543</td>
</tr>
<tr>
<td>M-SVMs + LEM (Quad)</td>
<td>78.4</td>
<td>0.73</td>
<td>0.64</td>
<td>0.61</td>
<td>74.8</td>
<td>0.541</td>
</tr>
<tr>
<td>M-SVMs + LEM (CE)</td>
<td>78.4</td>
<td>0.73</td>
<td>0.65</td>
<td>0.59</td>
<td>74.9</td>
<td>0.539</td>
</tr>
</tbody>
</table>

Table 1: Prediction accuracy of M-SVMs and combinations of M-SVMs used for protein secondary structure prediction. The first three machines are the ones of Weston and Watkins (WW), Crammer and Singer (CS), and Lee, Lin, and Wahba (LLW).

considered individually (trained on the whole training set), and in the framework of a combination (with a training set split into two or three according to the nature of the combiner) shows a gain in recognition rate induced by the combination which is always statistically significant with confidence exceeding 0.95. The LEMs appear slightly superior to the other combiners. They outperform the MLP according to the Sov’99 measure, which is promising in the perspective of the use of the corresponding predictions to provide constraints for tertiary structure prediction methods. Their cross-entropy is also inferior to that of the polytomous logistic regression.

6 Conclusions and ongoing research

In this article, a class of linear ensemble methods devoted to the combination, after an appropriate post-processing, of $Q$-category classifiers taking their values in $\mathbb{R}^Q$, has been introduced. Their specifications should make them well suited for the combination of M-SVMs. Indeed, the corresponding loss function can be chosen so that the outputs are class posterior probability estimates and their low sample complexity should prevent them from overfitting.

We are currently performing a large scale comparative study of their performance. The focus is laid on the quality of the probability estimates. In that respect, the baseline is provided by solutions implementing pairwise coupling (Bradley-Terry model) [11]. A touchstone consists in the use of the estimates to derive the emission probabilities of a hidden Markov model (see for instance [15]). In the specific case of protein secondary structure prediction, a successful post-processing of this kind should make it possible to bypass some of the limitations induced by the local approach (use of a sliding window). From a theoretical point of view, it should be
instructive to carry on this study by comparing the combinations of M-SVMs we consider with the extension of the LLW-M-SVM to a Bayesian model described in [27].

Acknowledgements The author would like to thank F. Thomarat for generating PSSMs for the sequences of the CB513 data set.

A Proofs of the main results of the article

A.1 Proof of Proposition 2

The proof of Proposition 2 is made up of three steps.

1. \( \forall v \in U_Q^N, Bv \in \mathbb{R}_+^Q \)

\[ \forall k \in [1, Q], \forall v \in U_Q^N, \beta_k^T v = \sum_{j=1}^N \sum_{l=1}^Q \beta_{kjl} v_{(j-1)Q+l}. \]  

Given the vector \( \beta_k \), the minimum of this inner product is obtained for a vector \( v \) satisfying:

\[ \forall j \in [1, N], \sum_{l: \beta_{kjl}=\beta_{kjl}^*} v_{(j-1)Q+l} = 1. \]

As a consequence, \( \min_{v \in U_Q^N} \beta_k^T v = \sum_{j=1}^N \beta_{kjl}^* \), from which it springs that \( \forall k \in [1, Q], \sum_{j=1}^N \beta_{kjl}^* \geq 0 \) is a necessary and sufficient condition of nonnegativity of the outputs.

2. \( \exists K : \forall v \in U_Q^N, 1_Q^T Bv = K \)

To derive the corresponding constraints, it suffices to notice that given any two vectors \( v^{(0)} \) and \( w \) in \( U_Q^N \), one can generate a finite sequence \( \{v^{(n)}\}_{1 \leq n \leq \infty} \) of vectors in \( U_Q^N \) such that \( v^{(n)} = w \) and \( v^{(n+1)} \) is deduced from \( v^{(n)} \) by applying an elementary step of the form:

(a) choose \( (j, l_1, l_2) \in [1, N] \times [1, Q] \times [1, Q] \) such that \( l_1 \neq l_2 \), \( v^{(n)}_{(j-1)Q+l_1} < 1 \) and \( v^{(n)}_{(j-1)Q+l_2} > 0 \);

(b) choose \( \delta \in \mathbb{R}_+^* \) satisfying \( v^{(n)}_{(j-1)Q+l_1} + \delta \leq 1 \) and \( v^{(n)}_{(j-1)Q+l_2} - \delta \geq 0 \);

(c) set \( v^{(n+1)} \) equal to \( v^{(n)} \) except for its components of indices \( (j-1)Q+l_1 \) and \( (j-1)Q+l_2 \) which are respectively set to \( v^{(n)}_{(j-1)Q+l_1} + \delta \) and to \( v^{(n)}_{(j-1)Q+l_2} - \delta \).

Given the generative algorithm detailed above, keeping the sum \( 1_Q^T Bv \) constant over the whole set \( U_Q^N \) boils down to ensuring that this sum does not vary when an elementary step is applied, so that

\[ \exists K \in \mathbb{R} : \forall v \in U_Q^N, 1_Q^T Bv = K \iff \forall (j, l_1, l_2) \in [1, N] \times [1, Q] \times [1, Q], \sum_{k=1}^Q (\beta_{kjl_1} - \beta_{kjl_2}) = 0 \iff \forall j \in [1, N], \forall l \in [1, Q-1], \sum_{k=1}^Q (\beta_{kjl} - \beta_{kjQ}) = 0. \]
3. $K = 1$

Once the conditions of the second step of the proof are satisfied, $1_Q^T B v$ does not depend on $v$ anymore. Thus, the constraint corresponding to $K = 1$ can be derived from an arbitrary choice of $v$ in $U_N^Q$. Setting $v = \frac{1}{\beta} 1_{NQ}$ gives:

$$K = 1 \iff 1_Q^T \sum_{k=1}^Q \sum_{j=1}^N \sum_{l=1}^Q \beta_{kjl} = 1 \iff 1_{NQ}^T \beta = Q.$$ 

The conjunction of this constraint and the ones obtained at the second step provides us with a stronger result that will prove useful in the sequel:

$$\forall l \in [1, Q], \quad \sum_{k=1}^Q \sum_{j=1}^N \beta_{kjl} = 1. \quad (9)$$

### A.2 Proof of Proposition 3

With Remark 1 in mind, to prove Proposition 3, it suffices to establish that for all vector $\beta$ satisfying the constraints of Proposition 2, one can exhibit a vector $\gamma$ in $\{ (U_N^Q)^{-1} \}^Q$ such that $\beta + \gamma \in \mathbb{R}^{NQ}_+$. For $\beta$ satisfying the constraints of Proposition 2 and $k \in [1, Q]$, let $I_k^-$ and $I_k^+$ be the subsets of $[1, N]$ such that $I_k^- = \{ j \in [1, N] : \beta_{kj} < 0 \}$ and $I_k^+ = [1, N] \setminus I_k^-$. The vector $\gamma^* = (\gamma^*)_{1 \leq k \leq Q}$ defined as follows:

$$\forall k \in [1, Q], \forall j_0 \in [1, N], \forall l \in [1, Q], \begin{cases} \text{if } j_0 \in I_k^-, \quad \gamma^*_{kj0l} = -\beta_{kj0} \\
\text{if } j_0 \in I_k^+, \quad \gamma^*_{kj0l} = \beta_{kj0} \sum_{j \in I_k^+} \beta_{kj} \end{cases}$$

with $\gamma^*_{kj0l}$ being the component of index $(j_0 - 1) Q + l$ of $\gamma_k^* \in \mathbb{R}^{NQ}$, meets the aforementioned requirements.

### A.3 Proof of Proposition 4

A possible construction of vectors $\Theta$ and $\theta$ corresponding to a given vector $\beta$ in $V_{NQ}$ is:

$$\forall j \in [1, N], \quad \Theta_j = \sum_{k=1}^Q \beta_{kjQ}$$

and

$$\forall (j, l) \in [1, N] \times [1, Q], \begin{cases} \text{if } \Theta_j > 0, \quad (\theta_{kjl})_{1 \leq k \leq Q} = \Theta_j^{-1} (\beta_{kjl})_{1 \leq k \leq Q} \\
\text{if } \Theta_j = 0, \quad (\theta_{kjl})_{1 \leq k \leq Q} = \frac{1}{\Theta_j} 1_Q \end{cases}.$$ 

Indeed, we deduce from (9) that

$$\sum_{j=1}^N \Theta_j = \sum_{j=1}^N \sum_{k=1}^Q \beta_{kjQ} = 1,$$

which implies that $\Theta \in U_N$. Proving that the vectors $(\theta_{kjl})_{1 \leq k \leq Q}$ belong to $U_Q$ is also straightforward. Conversely, if two vectors $\Theta$ and $\theta$ satisfy the hypotheses of Proposition 4, then obviously the vector $\beta$ of general term $\beta_{kjl} = \Theta_j \theta_{kjl}$ belongs to $V_{NQ}$.
A.4 Proof of Proposition 5

Given the hypotheses made in Section 2.1,  

\[
2 \mathbb{E}_{(X,Y)} [\ell_{\text{Quad}} (\tilde{g} (X), Y, \beta)] = \int_{X \times Y} \| t_y - g_{\beta} (x) \|_2^2 \, dP (x, y),
\]

and Fubini’s theorem for nonnegative measurable functions can be applied to the measure \( P \). This gives:

\[
2 \mathbb{E}_{(X,Y)} [\ell_{\text{Quad}} (\tilde{g} (X), Y, \beta)] = \mathbb{E}_X \left\{ \sum_{k=1}^Q \| t_k - g_{\beta} (X) \|_2^2 \, P (k|X) \right\}
\]

\[
= \mathbb{E}_X \left\{ \sum_{k=1}^Q \sum_{l=1}^Q \left[ \delta_{k,l}^2 - 2 \delta_{k,l} g_{\beta} (X) + g_{\beta} (X) \right] \, P (k|X) \right\}
\]

\[
= 1 - 2 \mathbb{E}_X \left\{ \sum_{k=1}^Q g_{\beta} (X) \, P (k|X) \right\} + \mathbb{E}_X \left\{ \sum_{k=1}^Q g_{\beta} (X)^2 \right\}
\]

\[
= 1 - \mathbb{E}_X \left\{ \sum_{k=1}^Q P (k|X)^2 \right\} + \mathbb{E}_X \left\{ \sum_{k=1}^Q [P (k|X) - g_{\beta} (X)]^2 \right\}.
\]

Thus,

\[
\arg \min_{\beta \in V_{N,Q}} \mathbb{E}_{(X,Y)} [\ell_{\text{Quad}} (\tilde{g} (X), Y, \beta)] = \arg \min_{\beta \in V_{N,Q}} \mathbb{E}_X \left\{ \sum_{k=1}^Q [P (k|X) - g_{\beta} (X)]^2 \right\}.
\]

Given (10), (4) simply expresses the consistency of the principle of empirical risk minimization for the class of functions \( \{\ell_{\text{Quad}} (\cdot, \cdot, \beta) : \beta \in V_{N,Q}\} \) (see Chapter 3 and 5 in [24]). Thus, to finish the proof of (4), it suffices to establish this consistency. Without going into particulars, this can be done by proving the finiteness of the capacity of the MLM. A result of this kind is provided by Theorem 1. The proof of (5) follows the same line of reasoning. For lack of space, it is omitted.

A.5 Proof of Lemma 1

The simplest way to establish Lemma 1 consists in making use of the alternative representation of the vectors of \( V_{N,Q} \) introduced in Proposition 4. Then, the proof is made up of two proofs by contradiction. Since they present no difficulty, we only give the sketch of the reasoning. The first proof by contradiction consists in establishing that if \( \beta \) is an extreme point of \( V_{N,Q} \), then necessarily \( \Theta \) is an extreme point of \( U_N \), which implies that

\[
\exists j_0 \in [1, N]: \begin{cases} 
\Theta_{j_0} = 1 \\
\forall j \in [1, N] \setminus \{j_0\}, \forall (k, l) \in [1, Q]^2, \beta_{kjl} = 0
\end{cases}
\]

The second proof by contradiction establishes that for all value of \( l \) in \([1, Q]\), the vector \((\theta_{kjl})_{1 \leq k \leq Q}\) must be an extreme point of \( U_Q \). This means that:

\[
\forall l \in [1, Q], \exists k_0 (l) \in [1, Q]: (\theta_{kjo(l)})_{1 \leq k \leq Q} = t_{k_0 (l)}.
\]

Combining the two partial results and using the fact that for all \( (k, l) \) in \([1, Q]^2\), \( \beta_{kjl} = \theta_{kjl} \), we get (6), which concludes the proof.
A.6 Proof of Proposition 6

Since Problem 3 is a linear programming problem in standard form, there exists an extreme point of the feasible region which is an optimal solution. Thus, (7) simply corresponds to a restriction of (6) to the extreme points of $V_{N,Q}$ that are also optimal solutions of Problem 3.

A.7 Proof of Theorem 1

The proof of Theorem 1 follows the sketch of the proof of Theorem 4.1 in [7]. In the same way as this proof, it is based on two lemmas.

Lemma 2 Let $\gamma \in (0, \frac{1}{2}]$ and $n \in \mathbb{N}$. If a subset $s_n = \{v_i : 1 \leq i \leq n\}$ of $U^N_Q$ is $N$-shattered with margin $\gamma$ by $\Delta G_3$, then there exists a subset $s_p$ of $s_n$ of cardinality $p$ equal to $\left\lfloor \frac{n}{2}\right\rfloor$ such that for every partition of $s_p$ into two subsets $s_{p,1}$ and $s_{p,2}$, the following bound holds true:

$$\left\| \sum_{v_i \in s_{p,1}} v_i - \sum_{v_i \in s_{p,2}} v_i \right\|_2 \geq 2 \left\lfloor \frac{n}{2}\right\rfloor^{-\gamma}.$$

Proof Suppose that $s_n = \{v_i : 1 \leq i \leq n\}$ is a subset of $U^N_Q$ $N$-shattered with margin $\gamma$ by $\Delta G_3$. Let $(s_n, c)$ witness this shattering. Without loss of generality, we can assume that $I(s_n)$ satisfies: $\forall i \in [1, n]$, $i_1(v_i) < i_2(v_i)$. According to the pigeonhole principle, there is at least one couple of indices $(k_1, k_2)$ with $1 \leq k_1 < k_2 \leq Q$ such that there are at least $p = \left\lfloor \frac{n}{2}\right\rfloor$ points in $s_n$ for which the couple $(i_1(v_i), i_2(v_i))$ is $(k_1, k_2)$. For the sake of simplicity, the points in $s_n$ are reordered so that the $p$ first of them exhibit this property. The corresponding subset of $s_n$ is denoted $s_p$. This means that for all vector $y = (y_i) \in \{-1, 1\}^n$, there is a function $g_\beta(y)$ in $G_3$ characterized by the vector $\beta(y) = (\beta_k(y))_{1 \leq k \leq Q} \in V_{N,Q}$ such that:

$$\forall i \in [1, p], \begin{cases} \text{if } y_i = 1, & \frac{1}{2} (\beta_{k_1}(y)^Tv_i - \max_{k \neq k_1} \beta_k(y)^Tv_i) - c_i \geq \gamma \\ \text{if } y_i = -1, & \frac{1}{2} (\beta_{k_2}(y)^Tv_i - \max_{k \neq k_2} \beta_k(y)^Tv_i) + c_i \geq \gamma \end{cases}$$

which implies that

$$\forall i \in [1, p], \begin{cases} \text{if } y_i = 1, & \frac{1}{2} (\beta_{k_1}(y)^Tv_i - \beta_{k_2}(y)^Tv_i) - c_i \geq \gamma \\ \text{if } y_i = -1, & \frac{1}{2} (\beta_{k_2}(y)^Tv_i - \beta_{k_1}(y)^Tv_i) + c_i \geq \gamma \end{cases}. \quad (11)$$

Consider now any partition of $s_p$ into two subsets $s_{p,1}$ and $s_{p,2}$. Consider any vector $y$ in $\{-1, 1\}^n$ such that $y_i = 1$ if $v_i \in s_{p,1}$ and $y_i = -1$ if $v_i \in s_{p,2}$. It results from (11) that:

$$\frac{1}{2} (\beta_{k_1}(y) - \beta_{k_2}(y))^T \left( \sum_{v_i \in s_{p,1}} v_i - \sum_{v_i \in s_{p,2}} v_i \right) - \sum_{v_i \in s_{p,1}} c_i + \sum_{v_i \in s_{p,2}} c_i \geq p\gamma. \quad (12)$$

Conversely, consider any vector $y$ such that $y_i = -1$ if $v_i \in s_{p,1}$ and $y_i = 1$ if $v_i \in s_{p,2}$. We have:

$$\frac{1}{2} (\beta_{k_2}(y) - \beta_{k_1}(y))^T \left( \sum_{v_i \in s_{p,1}} v_i - \sum_{v_i \in s_{p,2}} v_i \right) + \sum_{v_i \in s_{p,1}} c_i - \sum_{v_i \in s_{p,2}} c_i \geq p\gamma. \quad (13)$$
Combining (12), (13), and the Cauchy-Schwarz inequality, it appears that (whatever the sign of \( \sum_{i \in s_{n,1}} c_i - \sum_{i \in s_{n,2}} c_i \)) there is a function \( g_\beta \) in \( \mathcal{G}_\beta \) such that

\[
\frac{1}{2} \| \beta_k - \beta_l \|^2_2 \geq \gamma. \tag{14}
\]

Thus,

\[
\forall \beta \in V_{N,Q}, \forall (k, l) : 1 \leq k < l \leq Q, \quad \| \beta_k - \beta_l \|^2_2 \geq \left( \| \beta_k \|^2_2 + \| \beta_l \|^2_2 - 2 \beta_k T \beta_l \right) \leq \| \beta_k \|^2_2 + \| \beta_l \|^2_2.
\]

Since we know that \( \max_{\beta \in V_{N,Q}} \| \beta \|^2_2 = \sqrt{Q} \) (see Remark 2), we get

\[
\forall \beta \in V_{N,Q}, \max_{1 \leq k < l \leq Q} \| \beta_k - \beta_l \|^2_2 \leq \sqrt{Q}.
\]

A substitution of this upper bound in (14) then concludes the proof.

\[\Box\]

**Lemma 3** For all \( n \in \mathbb{N}^* \), all subset \( s_n = \{ v_i : 1 \leq i \leq n \} \) of \( U^N_Q \) can be partitioned into two subsets \( s_{n,1} \) and \( s_{n,2} \) satisfying

\[
\left\| \sum_{v_i \in s_{n,1}} v_i - \sum_{v_i \in s_{n,2}} v_i \right\|^2_2 \leq \sqrt{Nn}. \tag{15}
\]

**Proof** Let \( s_n = \{ v_i : 1 \leq i \leq n \} \subset U^N_Q \). Let \( \sigma = (\sigma_i)_{1 \leq i \leq n} \) be a Rademacher sequence, i.e., a sequence of i.i.d. random variables taking the values \(-1\) and \(1\) with probability \( \frac{1}{2} \). We have:

\[
E \left\| \sum_{i=1}^{n} \sigma_i v_i \right\|^2_2 = E \left[ \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_i \sigma_j v_i^T v_j \right] = \sum_{i=1}^{n} \sum_{j=1}^{n} v_i^T v_j E \left[ \sigma_i \sigma_j \right] = \sum_{i=1}^{n} \| v_i \|^2_2 \leq n \max_{v \in U^N_Q} \| v \|^2_2.
\]

Obviously, the vectors of \( U^N_Q \) whose \( \ell_2 \) norm is maximum are its vertices (extreme points). The corresponding value of the norm is \( \sqrt{N} \). Thus,

\[
E \left\| \sum_{i=1}^{n} \sigma_i v_i \right\|^2_2 \leq Nn.
\]

This implies that there exists a binary vector \( y = (y_i)_{1 \leq i \leq n} \in \{-1, 1\}^n \) such that

\[
\left\| \sum_{i=1}^{n} y_i v_i \right\|^2_2 \leq \sqrt{Nn}.
\]

Setting \( s_{n,1} = \{ v_i \in s_n : y_i = 1 \} \) and \( s_{n,2} = s_n \setminus s_{n,1} \) then concludes the proof.

\[\Box\]

With Lemmas 2 and 3 at hand, the proof of Theorem 1 is elementary.
Proof Let \( s_q = \{v_i : 1 \leq i \leq q\} \) be a subset of \( U_\mathbb{Q}^N \) N-shattered with margin \( \gamma \) by \( \Delta G_3 \). According to Lemma 2, there is at least a subset \( s_n \) of \( s_q \) of cardinality \( n \) equal to \( \left\lceil \frac{q}{Q} \right\rceil \) satisfying

\[
\left\| \sum_{v_i \in s_{n,1}} v_i - \sum_{v_i \in s_{n,2}} v_i \right\|_2 \geq \frac{2n}{\sqrt{Q}} \gamma
\]

for all its partitions into two subsets \( s_{n,1} \) and \( s_{n,2} \). Since, according to Lemma 3, there is at least one of these partitions for which (15) holds true,

\[
\frac{2n}{\sqrt{Q}} \gamma \leq \sqrt{Nn}
\]

which implies that

\[
q \leq \left( \frac{Q}{2} \right) \frac{NQ}{4\gamma^2},
\]

which is precisely (8).

References


Minimax choice of inspection program of fatigue-prone airplane structure

M. Hauka, Yu. Paramonov

Aviation Institute, Riga Technical University, Latvia
Email: maris.hauka@gmail.com
Aviation Institute, Riga Technical University, Latvia
Email: yuri.paramonov@gmail.com

Abstract: Development of the inspection program of fatigue-prone aircraft under limitation of airline fatigue failure intensity is discussed, using Markov Chains and Semi-Markov process theory. The minimax approach is offered for processing of results of approval full-scale fatigue test of airframe.

Keywords: Minimax, inspection program, Markov chains, reliability

1 Introduction

Limitation of fatigue failure probability (FFP) of fatigue-prone aircraft (AC) and fatigue failure rate (FFR) of airline (AL) is a problem of high priority. A lot of papers and books are devoted to this problem. In [1] the Markov Chains (MC) and Semi-Markov process with rewards (SMPW) theories are offered for its solution using exponential approximation of fatigue crack size growth function, $a(t) = \alpha \exp(Qt)$, where $\alpha$, $Q$ are parameters of fatigue crack trajectory (PFCT). The value $\alpha$ is called equivalent initial flow size (EIFS). (Note, it is not real initial flow size; it is only parameter of exponential approximation of fatigue crack trajectory!) The value $Q$ defines the speed of fatigue crack size growth in logarithm scale: $\log(a(t)) = \log(\alpha) + Qt$. PFCT are random variables. It is supposed that the cumulative distribution function (cdf) of the vector $(\alpha, Q)$ is known but some parameter of this cdf, $\theta$, is not known. Estimation of $\theta$ and the choice of inspection program under condition of limitation FFP up to specified life (AC retirement live), $t_{SL}$, or limitation of FFR of AL can be made using minimax processing of results of observation of some random fatigue cracks during AC type full-scale fatigue approval test. Specific feature of approval test is a decision to redesign the new AC type if some reliability requirements are not met. In [1] it was assumed that $\alpha$ is some constant. In this paper this assumption is eliminated.

2 Minimax choice of inspection program

Despite of all the simplicity, the equation $a(t) = \alpha \exp(Qt)$ gives us rather comprehensible result in the interval $(t_d, t_c)$, where $t_d$ is a time when the crack
becomes detectable \((a(t_d) = a_d)\) and \(t_c\) is a time when the crack reaches its critical size \((a(t_c) = a_c)\) and fatigue failure takes place. Corresponding random variables are defined by equations:

\[
T_d = \frac{(\log a_d - \log \alpha)}{Q} = C_d / Q, \quad T_c = \frac{(\log a_c - \log \alpha)}{Q} = C_c / Q.
\]

Let us denote \(X = \log Q\) and \(Y = \log C_c\), where \(C_c = \log a_c - \log \alpha\). From the analysis of the fatigue test data it can be assumed, that the log-log \(T_c = \log C_c / \log Q\) is distributed normally. It comes from the additive property of the normal distribution that it can takes place if either both \(\log C_c\) and \(\log Q\) are normally distributed or if one of these components is normally distributed while another one is constant. Contrary to [1] in this paper we consider the first case: vector \((X, Y) = (\log(Q), \log(C_c))\) has two dimensional normal distribution with vector-parameter \(\theta = (\mu_X, \mu_Y, \sigma_X, \sigma_Y, r)\). It is worth to note, that for the case when \(a_c\) and \(a_d\) are constants then cdf of \(C_d\) is completely defined by the distribution of \(C_c\) because \(C_d = C_c - \delta\), where \(\delta = \log(a_c / a_d)\).

For known \(\theta\) there are two decisions \(d_0\) and \(d_1\): the aircraft is good enough and the operation of this aircraft type can be allowed \((d_0)\) and redesign of aircraft should be made \((d_1)\). In the case of first decision the vector \(t = (t_1, ..., t_n)\), where \(t_i\) is the time moment of \(i\)-th inspection, should be defined also.

If \(\theta\) is known the different rules can be offered for the choice of structure of vector \(t\) : 1) every interval between inspections is equal to \(t_{SL} / (n + 1)\), 2) probability of failure in every interval is equal to \(P(T_c < t_{SL}) / (n + 1)\) ... In this paper we suppose that (just as in both mentioned examples) the vector \(t\) is defined by the fixed \(t_{SL}\) and choice of \(n\).

For substantiation of the choice of inspection number we should know FFP of AC and FFR and gain (GL) of AL as functions of \(n\). For this purpose the process of operation of AC can be considered as absorbing MC with \((n + 4)\) states. The states \(E_i, E_2, ..., E_{n+1}\) correspond to AC operation in time intervals \([t_0, t_1), [t_1, t_2), ..., [t_n, t_{SL})\), States \(E_{n+2}, E_{n+3}\), and \(E_{n+4}\) are absorbing states: AC is descarded from service when the SL is reached or fatigue failure (FF), or fatigue crack detection (CD) takes place.
Let in the transition probability matrix, $P_{AC}$, for corresponding process of AC operation the probability of crack detection during the inspection number $i$ be denoted as $V_i$; probability of failure in service time interval $t \in (t_{i-1}, t_i]$ as $q_i$, and probability of successful transition to the next state as $u_i = 1 - V_i - q_i$. In our model we also assume that an aircraft is discarded from service at $t_{SL}$ even if there are no cracks discovered by inspection at the time moment $t_{SL}$. This inspection at the end of $(n+1)$-th interval (in state $E_{n+1}$) does not change the reliability but it is made in order to know the state of aircraft (whether there is a fatigue crack or there is no fatigue crack). It can be shown that

$$u_i = P(T_d > t_i | T_d > t_{i-1}) = P(Q < C_d / t_i) / P(Q < C_d / t_{i-1}) = a_i / a_{i-1},$$

$$q_i = P(t_{i-1} < T_d < T_c < t_i | Td > t_{i-1}) =$$

$$= \begin{cases} 0, & \text{if } t_{i-1}C_c / C_d > t_i, \\ b_i / a_{i-1}, & \text{if } t_{i-1}C_c / C_d \leq t_i, \end{cases}, \quad i = 1, \ldots, n+1,$$

where

$$a_i = P(Q < C_d / t_i) = \int_{\ln \delta}^{+x} g_{ai}(y) d\Phi \left( \frac{y - \mu_y}{\sigma_y} \right),$$

$$g_{ai} = P(Q < C_d / t_i) = \Phi \left( \frac{\log \left( e^y - \delta \right) - \log t_i - \mu_{X_{ij}}} {\sigma_{X_{ij}}} \right),$$

$$b_i = P(C_c / t_i < Q < C_d / t_{i-1}),$$

$$= P(\log C_c - \log t_i \leq \log Q < \log(C_c - \delta) - \log t_{i-1}),$$

$$= \int_{\ln \delta}^{+x} g_{bi}(y) d\Phi \left( \frac{y - \mu_y}{\sigma_y} \right),$$

$$g_{bi}(y) = \max \left\{ 0, \Phi \left( \frac{\log \left( e^y - \delta \right) - \log t_{i-1} - \mu_{X_{ij}}} {\sigma_{X_{ij}}} \right), \right.$$
\[ \mu_{X/y} = \mu_X + r \frac{\sigma_X}{\sigma_Y} (y - \mu_Y), \]
\[ \sigma_{X/y} = \sigma_X \sqrt{1 - r^2}. \]

These probabilities can be calculated using Monte Carlo method also. The following equations can be used for modelling r.v. \( Y = \log C_e \sim N(\mu_Y, \sigma_Y^2) \)
and \( X = \log Q \sim N(\mu_X, \sigma_X^2) \) with some coefficient of correlation \( r \):
\[
Y = \eta_1 \sigma_Y + \mu_Y, \quad X = \eta_2 \sigma_X r + \eta_3 \sigma_X \sqrt{1 - r^2} + \mu_X,
\]
where r.v. \( \eta_1 \) and \( \eta_2 \) have standard normal distribution.

Recall that in matrix, \( P_{AC} \), there are three units in three last lines in matrix diagonal because states \( E_{n+2}, E_{n+3} \), and \( E_{n+4} \) are absorbing states: AC is discarded from service when the SL is reached or fatigue failure (FF), or fatigue crack detection (CD) take place.

In corresponding matrix for operation process of AL the states \( E_{n+2}, E_{n+3} \) and \( E_{n+4} \) are not absorbing but correspond to return of MC to state \( E_1 \) (AL operation returns to first interval). The other lines of \( P_{AC} \) and \( P_{AL} \) are the same.

For SMPW version of problem using \( P_{AL} \) we can get the airline gain
\[
g(n) = \sum_{i=1}^{n+4} \pi_i g_i (n), \quad \pi = (\pi_1, \ldots, \pi_{n+4}) \]
is the vector of stationary probabilities, which is defined by the equation system \( \pi P = \pi, \sum_{i=1}^{n+4} \pi_i = 1 \); AL operation rewards are defined in following way
\[
g_i (n) = \begin{cases} a_i \cdot u_i + b_i \cdot q_i + c_i \cdot v_i, & i = 1, \ldots, n+1, \\ d_i, & i = n + 2, \ldots, n + 4, \end{cases}
\]
where \( d_i \) is the reward related to successful transition from one operation interval to the following one and the cost of one inspection; \( b_i, c_i \) and \( d_i \) are related with transition to states \( E_{n+3} \) (FF), \( E_{n+4} \) (CD) and \( E_1 \). Let us note that if \( a = b = c = 1 \), \( d = 0 \) and time transition to state \( E_1 \) are equal to zero, then \( \pi_j = \pi_j g_j (n) / g(n) \) defines the part of time which SMP spends in state \( E_j, j = 1, \ldots, n+1 \); \( L_j = g(n) / \pi_j \) defines the mean return time for state \( E_j \).
Specifically, $L_{n+3}$ is the mean time between FF; so $\lambda_F = 1 / L_{n+3}$ is the FFR. It is worth to mention also, that the same value can be calculated and in another way. This value is equal to the ratio of aircraft failure probability, $p_F$, to the mean life of new aircraft, $L = g(n)/\pi_1$ (the mean time of renewal of AC (renewal operation of AL in the first interval)).

There are two versions of reliability requirements: A) limitation of FFR of AL; B) limitation of FFP of AC. More elaborately we consider case A. If $\theta$ is known we calculate the gain as function of $n, g(n,\theta)$, and choose the number $n_g$ corresponding to the maximum of gain: $n_g(\theta) = \arg\max_n g(n,\theta)$. Then we calculate FFR as function of $n, \lambda_F(n,\theta)$, and choose $n_\lambda$ in such a way that for all $n \geq n_\lambda$ the function $\lambda_\lambda(n,\theta)$ will be equal or less than some value $\lambda_{FD}$ (the “designed” FFR):

$$n_\lambda(\lambda_{FD},\theta) = \min\{n : \lambda_F(n,\theta) \leq \lambda_{FD}, \text{ for all } n \geq n_\lambda(\lambda_{FD},\theta)\}.$$ 

And finally we choose $n = n_{\lambda}(\lambda_{FD},\theta) = \max(n_g(\theta), n_\lambda(\lambda_{FD},\theta))$.

But we do not know $\theta$ and we can get only some estimate of this parameter, $\hat{\theta}$. Then, first of all, we should define some part of parameter space $\Theta_0$ in such a way that if $\hat{\theta} \notin \Theta_0$, then redesign of AC should be made.

If instead of $n_{\lambda}(\lambda_{FD},\theta)$ we use $\hat{n}_{\lambda}(\lambda_{FD},\theta)$ then real intensity FFR will be a function of random variable, $\lambda_F(\hat{n}_{\lambda},\theta)$. Let us define $\lambda_F(\hat{\theta},\lambda_{FD},\Theta_0) = \lambda_F(\hat{n}_{\lambda},\theta)$ if $\hat{\theta} \in \Theta_0$ and $\lambda_F(\hat{\theta},\lambda_{FD},\Theta_0) = 0$ if $\hat{\theta} \notin \Theta_0$ (service of this type of AC is not allowed). Corresponding expected value of FFR as function of $\theta$ has maximum because for “bad $\hat{\theta}$” we make redesign of airframe but for “very good $\hat{\theta}$” we do not need any inspection. Let us denote by $\lambda_{FD}(\Theta_0)$ the solution of the equation

$$\sup_{\theta} w_{\lambda}(\theta,\lambda_{FD},\Theta_0) = \lambda^*$$

(if the solution of this equation exist for specific $\Theta_0$),

where $w_{\lambda}(\theta,\lambda_{FD},\Theta_0) = E[\lambda_F(\hat{\theta},\lambda_{FD},\Theta_0)]$, $\lambda^*$ is the required FFR defined by specific aviation regulations. If after approval test we see that $\hat{\theta} \in \Theta_0$ then required inspection number $n = n(\lambda_{FD},\hat{\theta})$.

In similar way the choice of $n$ can be made for case B. Instead of
\[ w_j(\theta, \lambda_{FD}, \Theta_0) = \left\{ \int_{\Theta_0} \lambda_{FD}(n_{y^*}(\lambda_{FD}, x), \theta) dF_{\hat{\theta}}(x \mid \hat{\theta} \in \Theta_0) \right\} P(\hat{\theta} \in \Theta_0). \]

where \( F_{\hat{\theta}}(.) \) is cdf of \( \hat{\theta} \), the following p-set function \[1\] should be used
\[ w_p(\theta, p_{FD}, \Theta_0) = \sum_{i=1}^{n_p} P(t_{i-1}(\hat{\theta}) \leq T_d < T_e < t_i(\hat{\theta}), \hat{\theta} \in \Theta_0), \]
where 
\[ t_0 = 0, \ t_{n+1} = t_{SL}, \ t_i(\theta), i = 1, \ldots, n_{gg}, \ n_{gg} = \max(n_g, n_p(p_{FD}, \hat{\theta})), \ p_{FD} \]

is “designed” allowed FFP of AC which is used for the choice of \( n_p \) \( n_p(p_{FD}, \theta) = \min\{n: p_F(n, \theta) \leq p_{FD}, \ \text{for all } n \geq n_p(n_{FD}, \theta)\} \), function \( p_F(n, \theta) \) defines FFP of AC for specific \( n \) and \( \theta \).

### 3 Numerical example

In Table 2.1. of \[1\] some a priori information about fatigue crack growth function is given. Using this information for \( \alpha_c = 237.8 \) mm and \( \alpha_d = 20 \) mm (these values already have been used in \[1\]) we get the following estimate of parameter \( \hat{\theta} = (\hat{\mu}_x, \hat{\mu}_y, \hat{\sigma}_x, \hat{\sigma}_y, \hat{f}) = (-8.587, 1.942, 0.155, 0.0779, 0.796) \), see Table 1.

<table>
<thead>
<tr>
<th>Crack #</th>
<th>( \log(\alpha) )</th>
<th>( Q )</th>
<th>( X = \log(Q) )</th>
<th>( Y = \log(Cc) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>-1.2513</td>
<td>0.000186</td>
<td>-8.58976388</td>
<td>1.9055189</td>
</tr>
<tr>
<td>77</td>
<td>-1.8768</td>
<td>0.000195</td>
<td>-8.542511</td>
<td>1.9944819</td>
</tr>
<tr>
<td>78</td>
<td>-1.2445</td>
<td>0.000161</td>
<td>-8.73410619</td>
<td>1.9045069</td>
</tr>
<tr>
<td>92</td>
<td>-1.697</td>
<td>0.00022</td>
<td>-8.42188301</td>
<td>1.9697097</td>
</tr>
<tr>
<td>93</td>
<td>-1.5102</td>
<td>0.000207</td>
<td>-8.48279176</td>
<td>1.9433016</td>
</tr>
<tr>
<td>112</td>
<td>-2.5329</td>
<td>0.000228</td>
<td>-8.38616493</td>
<td>2.0800033</td>
</tr>
<tr>
<td>116</td>
<td>-0.6479</td>
<td>0.000154</td>
<td>-8.77855796</td>
<td>1.8114796</td>
</tr>
<tr>
<td>129</td>
<td>-1.4226</td>
<td>0.000157</td>
<td>-8.75926475</td>
<td>1.9306798</td>
</tr>
<tr>
<td>Mean</td>
<td>-1.5229</td>
<td>0.0001885</td>
<td>-8.58688044</td>
<td>1.9424608</td>
</tr>
<tr>
<td>StdDev.</td>
<td>0.548084</td>
<td>2.898778E-05</td>
<td>0.155128668</td>
<td>0.0778895</td>
</tr>
<tr>
<td>Corr.coeff. of X and Y</td>
<td>0.796437</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is supposed that all inspection intervals are equal. The following definition of components of AL income are used: for all \( i = 1, \ldots, n+1 \)
\[ a_i = a(n) = a_0(n) + d_{\text{sum}t_{SL}}, \ \text{where} \ a_0(n) = a_{00}t_{SL} / (n+1) \]
the reward, related to successful transition from one operation interval to the following one, \( a_{01} \) defines the reward of operation in one time unit (one hour or one flight); \( d_{\text{insp}}t_{SL} \) is the cost of one inspection (negative value) which is supposed to be proportional to \( t_{SL} \); \( b_i = b_{01}t_{SL} \) is related to FF (negative value), \( c_i = c_{01}a_{01}(n) \) is the reward related to transitions from any state \( E_1, \ldots, E_{n+1} \) to the state \( E_{n+4} \) (it is supposed to be proportional to \( a_0 \) because it is a part of \( a_0 \)); \( d_i = d_{01}t_{SL} \) is negative reward, the absolute value of which is the cost of new aircraft acquisition after events SL, FF or CD and transition to \( E_i \) take place. In numerical example we have used the following values: \( b_{01} = -3 \); \( d_{\text{insp}} = -0.05 \); \( a_{01} = 1 \); \( c_{01} = 0.05 \); \( d_{01} = -0.3 \). The set \( \Theta_0 \) is defined in following way: redesign of the AC type should be made if estimate of mean AC life is small (\( T_e < t_{SL} \)) or speed of fatigue crack growth is large (\( \log Q > \hat{\mu}_X + \sigma_X \)). For \( \lambda_{FD} = 0.0000001 \) in Fig. 1 (left) results of calculation of \( w_{\lambda}(\theta, \lambda_{FD}, \Theta_0) \) and corresponding (right) \[ w_{\rho_\lambda}(\theta, \lambda_{FD}, \Theta_0) = \sum_{i=1}^{n+1} P(t_{i-1}(\hat{\theta})) < T_d < T_e < t_i(\hat{\theta}), \hat{\theta} \in \Theta_0) \), where \[ t_0 = 0, \quad t_{n+1} = t_{SL}, \quad t_i(\theta), i = 1, \ldots, \hat{n}_{g\lambda}, \hat{n}_{e\lambda} = n_{g\lambda}(\lambda_{FD}, \hat{\theta}) \], as the function of \( \mu_X \) for \( (\mu_{Y_1}, \ldots, \mu_{Y_5}) = (1.55, 1.75, 1.94, 2.14, 2.33) \) in vicinity of its maximum are shown. It was supposed that the vector \((\sigma_X, \sigma_Y, r)\) is the same for different vectors \((\mu_X, \mu_Y)\) and it is equal to test estimate \((0.155128668, 0.0778895, 0.796437)\). (Recall, \( (\mu_Y - \mu_X) \) is equal to the \( E(\log(T_e)) \)).

![Fig. 1.](image)

Fig. 1. The \( w_{\lambda}(\theta, \lambda_{FD}, \Theta_0) \) and \( w_{\rho_\lambda}(\theta, p_{FD}, \Theta_0) \) as function of \( \mu_X \)
Maximum value of $w_j(\theta, \lambda_{FD}, \Theta_0)$ is equal to $1.523 \times 10^{-7}$. Maximum value of $w_p(\theta, \lambda_{FD}, \Theta_0)$ is equal to 0.000677. Suppose that these values satisfy the requirements: required FFR of AL and required FFP of AC. And now let us suppose that in real test we have got $\hat{\mu}_X = -8.5885$, $\hat{\mu}_Y = 1.942460769$, see Fig. 2. (These values was considered already in [1]). After calculations FFR and airline gain, see Fig. 3, for these specific parameters we find required number of inspections: $n = \max(n_g, n_\lambda) = \max(3, 4) = 4$. It apperas that the influence of scatter of EIFS, $\alpha$, apperas to be very significant. After similar calculations of $w_j(\theta, \lambda_{FD}, \Theta_0)$ as function of $\mu_X$ for $\sigma_Y = 0.00001$, $r = 0$ we get its maximum value equal to $1.87 \times 10^{-8}$. It is nearly 10 times lower than in previous case (when $\sigma_Y = 0.0778895$). So it is very important to take into account the scatter of EIFS, $\alpha$.

![Fig. 2. Example of fatigue crack size as function of flight number](image1)

Fig. 2. Example of fatigue crack size as function of flight number

![Fig. 3. Airline gain and FFR as functions of inspection number for specific $\hat{\mu}_X = -8.5885$, $\hat{\mu}_Y = 1.942460769$.](image2)

Fig. 3. Airline gain and FFR as functions of inspection number for specific $\hat{\mu}_X = -8.5885$, $\hat{\mu}_Y = 1.942460769$.

References

APPLICATION OF A COMBINED SHEWHART-CUSUM CHART FOR BINOMIAL DATA

Elisa Henning, Robert Wayne Samohyl, Olga Maria Formigoni Carvalho Walter, Andrea Cristina Konrath and Custodio da Cunha Alves

University of Santa Catarina State, Department of Mathematics, Brazil
Email: dma2eh@joinville.udesc.br
Federal University of Santa Catarina, Department of Production Engineering, Brazil
Email: samohyl@yahoo.com
Federal University of Santa Catarina, Department of Production Engineering, Brazil
Email: olgaformigoni@gmail.com
Federal University of Santa Catarina, Department of Computer Science and Statistics, Brazil
Email: andreack@inf.ufsc.br
University of Joinville Region, Brazil
Email: custodio.alves@univille.net

Abstract: The traditional Shewhart control charts are considered effective in detecting large changes in mean, variance or the fraction nonconforming, while cumulative sum control charts (CUSUM) are recommended for recognizing small and moderate changes in these parameters. Neither of the charts mentioned will perform well in all situations. One possible solution to this problem is to combine multiple charts to cover changes of various magnitudes. Thus, a combined Shewhart-CUSUM chart is intended to increase the sensitivity of the CUSUM procedure to large shifts. This work brings some contributions to the development and improvement of combined Shewhart-CUSUM charts for binomial data. Thinking of applications, a methodology was developed for building a combined scheme including analysis of the assumptions required (goodness of fit, autocorrelation and overdispersion). This methodology was applied to real data. The results of the applications were satisfactory, thus validating the methodology developed.

Keywords: Combined Shewhart-CUSUM chart, Binomial data, Application

1 Introduction

The traditional Shewhart charts are the most widely known and applied. The simplicity of the decision rule conditioned only on analyzing the last point signaled on the chart facilitates evaluating if such a point signals a special cause. However, the major disadvantage is utilizing only the information about the process contained in the last point and ignoring any information given by the entire sequence of points. This characteristic makes these control charts insensitive to slight changes in the process, on the order of 1.5 standard deviations or less (MONTGOMERY, 2004).

Although extremely efficient, these charts are not the only tools available. In some cases, other types of control charts may advantageously complement or substitute the traditional Shewhart chart. Such is the case of the CUSUM (Cumulative Sum) and the EWMA (Exponentially Weighted Moving Average) control charts. These charts are good to monitor processes subject to small shifts, whose decision
concerning the statistical control state is based on information accumulated from previous samples and not merely the last of them. With this, it is possible to signal small changes more quickly, as well as identify the moment in time in which a change in the process occurs. However, if the magnitude of the change were unknown or if there were an alteration over time, none of the previously-mentioned control charts would offer adequate performance concerning small and large alterations. But, it is possible to combine multiple charts in order to monitor diverse magnitudes of change, adding Shewhart control limits to CUSUM charts in order to detect both small and large shifts (TSUNG; WANG, 2010).

2 Combined Shewhart-CUSUM chart for binomial data

The unilateral combined Shewhart-CUSUM chart for binomial data incorporates a Shewhart np control chart for the number of nonconforming units and a binomial CUSUM chart on the same axis. The Shewhart np control chart is extensively utilized to monitor processes which produce a certain percentage of defective processes (DURAN; ALBIN, 2009; COSTA; EPPRECHT; CARPINETTI, 2005; SAMOHYL, 2009). It is popular in factories in which the use of the statistical process control is in its initial implantation phase (SAMOHYL, 2009). On the other hand, a binomial CUSUM chart examines the number of nonconformities accumulated in a sequence of samples. Its objective is to detect either increases or decreases in the number of nonconformities. A unilateral binomial CUSUM chart may be applied to detect an increase in the expected value of nonconforming items, from the nominal value \( np_0 \) to \( np_1 \), where \( n \) corresponds to the sample size. \( p_0 \) is the proportion of nonconforming items considered to be within statistic control, while \( p_1 \) is the proportion outside statistical control that one wishes to detect (HAWKINS; OLWELL, 1998; BOURKE, 2001; WU; JIAO; LIU, 2008).

Thus, if \( Y_i \) were a series of independent samples with \( i \in \mathbb{Z}^+ \) of size \( n \) and binomial distribution, the combined Shewhart-CUSUM binomial chart is obtained through plotting the statistics \( C_i \) and \( X_i \), with respect to the sample \( Y_i \),

\[
C_i = \max(0, C_{i-1} + X_i - k),
\]

where \( C_i \) is the CUSUM statistic with \( C_0 = u, 0 \leq u \leq h \); \( X_i \) is the number of nonconforming items in the sample \( Y_i \) with \( i = 1, 2, \ldots \); \( k \) is the CUSUM reference value and depends upon the magnitude of the change one desires to detect, and \( h \) is the CUSUM control limit. The Upper Control Limit (UCL) from the conventional Shewhart np-chart is expressed as

\[
UCL = np + 3\sqrt{np_0(1 - p_0)}.
\]

These limits are obtained through an approximation to the normal distribution for the value of the standard error. The use of the approximation is standard procedure.
in this area, however it is suggested that exact limits be calculated utilizing the
binomial distribution, if the situation needs it.

The two lines \( C_i \) and \( X_i \) and the two control limits, \( h \) and UCL (Shewhart Upper
Control Limit), are then plotted. If \( C_i \) goes beyond the superior control limit \( h \) or
\( X_i \) goes beyond the UCL, this signifies that the process is not in statistical control.
The reference value \( k \) from the CUSUM charts for counts is determined by the rate
of acceptable counts \( (np_0) \) and the rate of counts that one wishes to detect \( (np_1) \).
Gan (1993) proved that for a binomial CUSUM, the \( k \) value obtained through the
equation below may be considered an optimal value in the ARL sense in detecting
an upward shift of magnitude in the parameter \( p \) (MOR AIS; PACHECO, 2006).

Average Run Length (ARL) corresponds to the average value of the number of
observations which must be plotted to indicate a condition outside of statistical control. ARL\(_0\) indicates the average number of samples collected until the emission of a signal during the period under control. The ARL\(_1\) represents the average number of samples collected until the emission of a signal that indicates a situation that is out of control.

The upper limit \( h \) of the CUSUM is determined as a function of ARL\(_0\) and ARL\(_1\).
The precise relationship among the three parameters is not straightforward. There are various procedures used in the literature to calculate ARL\(_0\) and ARL\(_1\) for a binomial CUSUM. The most common is as a Markov process (BROOK; EVANS, 1972; HAWKINS, 1992; REYNOLDS; STOUMBOS, 1999). Other procedures are based on Wald’s (1947) approximations, Siegmund’s (1985) proposal, and the work of Reynolds and Stoumbos (1999, 2001).

Performance measurements for combined charts can be calculated in several ways already well documented in the literature. Markov chains were applied by Lucas (1982); Yashchin (1985) and Morais and Pacheco (2006) to continuous and discrete variables with a Poisson distribution, and simulation procedures by Rocha (2004) for continuous variables. Montgomery (2004), Souza and Samohyl (2008), Coelho (2008) and Wu, Jiao and Liu (2008) all approached the combination of control charts while considering the probability of false alarms.

In general, if there are 2 control charts monitoring the same variable and if each chart is characterized by the same probability of type I error, then the combined probability (type I error) for the combined control charts is \( \alpha = 1 - (1 - \alpha)^2 \), where \( (1 - \alpha)^2 \) is the probability that the two variables are simultaneously within their control limits. Thus, when two different charts monitor the same variable, the false alarm rate of the combined chart \( \alpha_{cs} \) will be a combination of the individual rates of each, \( \alpha_s \) and \( \alpha_c \) for Shewhart (S) and CUSUM (C), respectively. An expression which summarizes this combination is given by \( \alpha_{cs} = \alpha_s + \alpha_c - \alpha_s \alpha_c \).

As such, two control charts with individual false alarm rates of 1% for example generate a combined chart with a false alarm rate of nearly 2%. Looking to maintain the combined rate at its original value of 1%, the control limits should be recalculated, resulting in values which are at least as distant from the center line as
before and, if appropriate, further from it (COELHO, 2008; SOUZA; SAMOHYL, 2008).

3 Planning a Shewhart-CUSUM chart for binomial data

This section describes some procedures for planning a binomial combined Shewhart-CUSUM chart (HENNING, 2011). This sequence of steps takes into account the assumptions necessary for the application of the chart. Then we have an application involving the control of quality in the labeling of bottles. As a complement, the combined chart is compared with Shewhart and CUSUM individual procedures.

Set a series of $N$ samples of size $n$ with $p_0$. To construct the chart, it is recommended to:

i. Define the ratio $p_1$ that we want to detect;
ii. Define $\text{ARL}_0$ (or the false alarm rate) of the combined chart;
iii. Check the following assumptions:
   a. Adherence to the binomial distributions;
   b. Absence of overdispersion;
   c. Independence;
iv. Calculate the $\text{ARL}_0$ values for the individuals charts;
v. Calculate the reference value $k$;
vii. Approximate the $h$ of the CUSUM;
vii. Calculate the UCL of the Shewhart “np”;
viii. Verify if $\text{UCL} \leq h + k$, otherwise, it is not a combined chart.

When not checked for adherence to the binomial distribution, it is not recommended to apply the chart. The CUSUM is very sensitive to the removal of the distribution (OLWELL; HAWKINS, 1998). If there is overdispersion, we recommend its removal. This can be done by applying the chart to the residuals of a generalized linear model, such as a binomial model. But, so far there is no record of papers covering this kind of application.

In the case of dependence, sequential in time, as is the case for most applications in statistical process control, the absence of autocorrelation should be verified. If there is autocorrelation, techniques that remove or capture the autocorrelation should be used. Although there are several works for Shewhart and CUSUM charts for autocorrelated discrete data, there are still no studies covering combined charts.

4 Application

A combined chart is applied to data adapted from Devore (2006). The data set consists of a sample of $n = 1000$ items, selected every 25 consecutive days. These data are used in phase I, when the control limits are defined. The proportion $p_0$ is 0.0608 and $p_1 = 0.073$, an increase of 20% in $p_0$. The $\text{ARL}_0 = 370$ corresponds to a false alarm rate of 0.27%.
Thus a combined chart was planned with \( k = 6.7; \ h = 17.1 \) and UCL = 15. After verification of the assumptions, for the phase II of monitoring, a situation out of statistical control was simulated, with \( p_1 = 0.073 \) and a change purposefully large, about 180\%, in the 30th sample (5th sample after the change in \( p_0 \)). This ensured that the average of the simulated data was equal to \( np_1 \). The combined chart was applied to the data and the result is shown in Figure 1. The combined chart signals both changes, the change in the 30th sample, by the Shewhart part, and the CUSUM shows the increase in the average, with \( h \) exceeded at the 44th sample.

![Combined Shewhart-CUSUM Chart](image)

**Fig.1. Combined Shewhart-CUSUM Chart applied to the data.**

Then the results of the combined charts are compared with individual CUSUM and Shewhart charts. The CUSUM has parameters \( k = 6.7 \) and \( h = 16.5 \). The Shewhart Upper Control Limit is UCL = 13.25. The CUSUM chart is shown in Figure 2 and the Shewhart chart in Figure 3.

The CUSUM chart, Figure 2, indicates the increase in mean \( np_0 \), signaled at the 44th sample. The CUSUM, however, does not signal the sudden change in the 30th sample. The Shewhart chart, Figure 3, indicates only the change in the 30th sample, but cannot detect the increase in average. This is an example in which the combined chart is more effective than the individual charts.

We know that a single application is not enough to validate this method. More examples with simulated and real data are shown in Henning (2011).
Fig. 2. CUSUM Chart applied to the data.

Fig. 3. Shewhart Chart applied to the data.
5 Conclusions and Final Considerations

This paper dealt with combined Shewhart-CUSUM charts for attributes, specifically np charts in which the underlying distribution is binomial. Some steps to implement the chart were laid out. In the situation proposed in the application, the combined chart signaled both major and small shifts, that were not be identified by the individual charts. We can also extend the proposed methodology for combined charts for Poisson data, taking into account the specificities of this.

References

THE METHOD OF EXPERT EVALUATIONS
IN BANKNOTE QUALITY ASSESSMENT

Victor Ionov
Independent Expert-Consultant (Russia)
Member of Statistics and Research Committee
The International Association of Currency Affairs (IACA)

Dear Mr. Chairman, Ladies and Gentlemen!

I'd like to say about some developments in banknote quality and security assessment. My presentation includes 4 main subjects:

1. The banknote as an object of investigation
2. What is “banknote quality and security”?
3. How to evaluate a banknote
4. Case studies and some results of banknote assessment

Let me start from brief introduction.

First of all I'll say few words about IACA (The International Association of Currency Affairs) as non-profit but high professional association. Currently IACA members are Issuing Authorities and Central Banks around the world, International Security Authorities, Currency Producers, Currency Users, Volume Currency Handlers and Industry Suppliers as well as other companies and individual members closely and professionally connected with cash circulation.

Seems to be there is no necessity to accent your attention on importance of this problem. Anyway, secondly I mention 5 main drivers which are putting forward the problem we’re discussing here. Those are:

- permanent growth of cash in circulation both in Europe and worldwide,
- growth and high level of counterfeiting,
- intensive development and introduction of new banknotes and security features,
- importance of banknote public perception,
- IT progress and competition between cash and other payment instruments/

And finally I’d like to say something about how those drivers moved me to make this case study. Including some references to my presentations, articles and the books.

1. The banknote as an object of investigation

It's traditionally accepted to understand as «quality» the aggregate of objectively existing properties and characteristics which level is stipulated by parameters defining value in use of product. The banknote, as such, evidently, is a mass product of industry. However, at the same time, I would like to name a banknote the work of art as well.

Actually, let's compare two objects: a banknote and a picture. Which one is more complicated and is saturated with different properties and characteristics?

Let's start with a substrate: for a picture it is canvas, cardboard or paper. If we compare that one with banknote substrate we can see that besides of special paper with multitone («picture in
the picture») and other kind of watermarks, it includes also security fibers, confetti, security threads, etc.

Further, appearance: the banknote is, practically, the engraving of a portrait or other image. That is the product incomparable with some creations of cubism artists, for example (See Figure 1), due to the difficulty of reproduction (I don't want to criticize anybody). I even don't speak about Kazimir Malevich well-known picture «Black Quadrate». We can add here for consideration every possible ornament, vignettes, grids and other special patterns in banknote.

Fig.1. Picasso «The Muse» and De La Rue Giori «Le Mont 2000»

And, at last, if to compare a variety of inks and the methods of printing used in banknote production with those paints and methods of drawing which are applied in painting – the advantage, seems to be, will be on the first party.

Besides that, to the painter some errors and distortions of proportions are pardonable, whereas in banknote production any modifications of inks or figures are inadmissible. Just add to this, that pictures (even sketches) cannot be combined in halves, crumpled, taken by dirty hands, dropped to the pools, and also laundered together with shirt or trousers. The pictures aren't counted a lot of times by counters, or sorters, or other devices working with banknotes. The experts, who are checking picture authenticity, use the advanced devices and work rather slowly.

Thus, we have right to tell, that banknote is a complicated work of art in industrial fulfillment with application of state-of-art technologies. The same words regards to the most of high-security printing products. However, these preliminary notes are made just to accent your attention on art aspect of banknote quality.

2. What is “banknote quality and security”?  

Very often the concept of «quality» is considered in two various aspects: as correspondence of an item to some standard (for example, all metal nuts should have the identical sizes, the thread, quality of metal, etc.) though the standard can be primitive enough. The second aspect is quality, as aggregate of useful features, properties and characteristics, defining ability of object to satisfy the stipulated or prospective requirements.

In our case we shall speak about quality of banknotes, as aggregate of their characteristics, such as wear resistance, limpness, stability to soiling, folding, rupture and tearing, art dignities and color scale, convenience of counting and identification, set of security elements, number of used kinds of printing, steadiness against counterfeiting and other. And “security” (or protection) of banknote as its ability to resist to counterfeiting is one of the most important part of banknote quality. In both cases we'll keep in mind the cost of production certainly.

The choice of quality performances on which the estimation is carried out should precede the assessment of production quality level. The part of them can be measured; however the significant part of banknote characteristics is very difficult to define quantitatively and accordingly to evaluate their quality. And the most of security features are the closest to this second part. So we need “to digitalize” our feel of paper, raised printing, design and layout of banknote, optically variable features like OVD, OVI etc. As European Central Bank (ECB) tells us: “Feel, Look, Tilt, Check”.
3. How to evaluate a banknote

To do that we need to develop methodology of assessment including the list of estimated parameters and necessary procedures. Seems to be the best and possibly the only method to get quantitative estimation of quality indicators is the method of expert evaluations.

I think there is no necessity to remind the theory of this method’s use. In practice its application can be viewed as described below.

At the first stage the customer (State Authority, Central Bank, Printworks, International Conference, etc) formulates the purposes of case study (assessment of existing banknotes or new series development, competitive analysis, creation of standardized parameters system), establishes Expert Commission including Expert Group(s) and Workgroup for work organization and processing results of experts' judgements and evaluations.

Seems to be, the Commission structure should include 3 Expert Groups: banking experts/cashiers, security printers and experts-criminalists. That's reasonable also to engage other target groups (shop cashiers, retailer, older persons, etc.).

At the second stage Workgroup makes the choice of methods, ways and procedures of estimation. Main features of estimated product (substrate, inks, OVD, security thread, printing methods, etc.), work timetable, labor input, and possible use of obtained results have to be taken into account.

Definition of the operations list which should be executed by experts is carried out by Workgroup proceeding from the estimation purposes, available information, the chosen estimation methods, ways and procedures, terms and conditions of estimation process. To get experts evaluations various methods (group or individual) and procedures of inquiry (interviewing, questioning or the mixed questioning) are used.

For realization of experts inquiry by Questioning method the Workgroup makes the questionnaire which consists of an explanatory note and inquiry map. In the explanatory note the purpose and procedure of questioning are stated; with specific example it's shown, how to fill in the columns of inquiry map; requests to banknote and classification groups of parameters and other information are described if needed. Inquiry map includes properly organized list of the questions, for each of those the certain answer style is provided.

The third stage of work is carried out by Expert Group(s) which members express their evaluations and judgements according to the methods and procedures established at the second stage.

At the fourth, final stage the members of Workgroup are processing expert evaluations and the Experts' Commission Report is made out.

As result of the second stage of work the Commission gets the set of parameters describing hard-formalizable properties of banknote (for example, kinds of watermarks used at manufacturing, the types of security thread, inks and other). Each parameter can be separated, in turn, to its components.

The estimation of each component is executed by experts at the third stage under established criteria. For example, by five-point scale with tenth shares (the maximum mark is 5,0). Furthermore experts determine the conditional factor (importance or «weight») of each component in the 100-th shares (the sum should be equal to 1) because the most frequently used method in expert evaluation is the method of the weighted average parameter.

As results of the third stage the Workgroup receives filled in questionnaires in which the points of each parameter are shown and the factors of importance («weight») of this parameter are determined.
At fourth stage the Workgroup processes the data of experts’ inquiry, calculating the Weighted Average Quality Index (WAQI) of banknote $Q_j$ for each $j$-th expert:

$$Q_j = \sum_{n=1}^{N} m_{nj} q_{nj}$$

Where $m_{nj}$ - weight factor of $n$-th parameter, and $\sum_{n=1}^{N} m_{nj} = 1$ ;

$q_{nj}$ - the point of $j$-th expert given to $n$-th parameter;

$N$ - number of estimated parameters.

General Weighted Average Quality Index $Q$ as equidistributed estimate of experts’ evaluations is calculated further:

$$Q = \frac{\sum_{j=1}^{J} Q_j}{J}$$

Where $J$ is the number of experts.

In case of several ($L$) Expert Groups, the formula becomes insignificantly complicated. However, in this case it is expedient to use weight factors of each group evaluations. So:

$$Q = \sum_{l=1}^{L} m_l Q_l$$

Or, with allowance from previous formula

$$Q = \sum_{l=1}^{L} \frac{\sum_{j=1}^{J} Q_j}{J}$$

This index will characterize the banknote quality taking into account evaluations of experts and Expert Groups.

It is possible to apply various complex cost indexes to get financial and economic estimation of quality of existing series of banknotes, elementary of which are Adjusted Present Value of banknote production (the ratio of production cost to the Weighted Average Quality Index) and Specific Quality Banknote Index (the ratio of the Weighted Average Quality Index to the nominal value of a banknote in US dollars or Euro, for example).


Due to results of tests and expert evaluations not only rating parameters of banknote counters, but also the performances most essential from the point of view of the majority of cashiers had been determined.

The same method (simplified) was applied in the middle of 1998 - during realization of questioning after results of denomination in Russia. 240 cashiers from 25 various organizations and target groups were in 5 Expert Groups and filled in their judgments to questionnaire for Goznak Research Institute.
4. Case studies and some results of banknote assessment

In January, 2003 (and later this year) due to kind assistance from some leading specialists experts' quality evaluation of three samples («Banknote 2000 High-Sec» of Giesecke and Devrient, «Le Mont 2000» of De La Rue Giori and «Turgenev 2000» of Goznak) and four real banknotes which are in circulation (5 US dollars, 5 euro, 100 Russian rubles and 100 Slovak krones) has been conducted.

Denominations of banknotes have been chosen proceeding from reasons approximate conformity of nominal cost in Russian rubles at the exchange rate. The samples were investigated just as examples of high quality.

Fifteen high-qualified specialists have taken part in that work (Forensic Science Center of the Ministry of Internal Affairs, Tver Branch of Bank of Russia and cashiers from five leading commercial banks). Thus, experts represented 3 various target groups distinguished as under the approach to estimation, working experience and on a level of technical equipment: 3 forensic experts, 4 experts from Bank of Russia and 8 cashiers from commercial banks.

In a role of the chairman of Expert Commission (and simultaneously whole working group) the author of this speech acted (by whom the technique of banknote quality estimation has been developed as well). During their work and fillings of questionnaire experts evaluated banknotes and samples on chosen parameters describing quality of paper, the applied inks and types of printing, various security elements (see below, Table 1, the first column). The 5-point scale has been applied with the equidistant gradation.

Results of expert judgments were processed. According to complex method of valuation Weighted Average Quality Index (WAQI) for each expert on the basis of which the rating for each researched category (samples / banknotes) was defined. The general average result of estimations and ratings has been further determined accordingly to all 15 experts as a whole and for each Expert Group.

The times has changed so new demands has appeared. And in October 2004 we have done new, 2nd round of banknote quality assessment. The style and details of estimation tests were upgraded a bit, keeping in mind previous experience and new tasks arisen. In this case experts estimated 3 samples and 5 real banknotes which are in circulation (NexGen 20 US dollars, 10 euro, 200 Slovak krones, modified 500 and 1000 Russian rubles. So we went to the next denomination and protection level. Primarily 4 denominations were planned to estimate and higher denomination banknote 1000 RUR was added at the request of Bank of Russia.

The most of experts participated in previous round of banknote quality estimation, so they knew approach, estimated characteristics and the style of the testing very well. It’s saved the time and highly improves the assessment process. Their recommendations were included into procedures of course. The results of this 2nd round of banknote quality assessment provide possibility to determine effectiveness of introduction of different new security features from the point of view of different expert target groups and in a whole.

<table>
<thead>
<tr>
<th>Quality Index</th>
<th>Importance Mark</th>
<th>G&amp;D BN2005</th>
<th>DLRG LeMont</th>
<th>Goznak Tolstoy</th>
<th>USD 20</th>
<th>EUR 10</th>
<th>SKK 200</th>
<th>RUR 500</th>
<th>RUR 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Quality</td>
<td>4,69</td>
<td>4,62</td>
<td>4,20</td>
<td>4,41</td>
<td>4,21</td>
<td>4,01</td>
<td>4,59</td>
<td>4,37</td>
<td>4,38</td>
</tr>
<tr>
<td>Size of Banknote</td>
<td>3,43</td>
<td>3,29</td>
<td>3,71</td>
<td>2,39</td>
<td>4,44</td>
<td>3,94</td>
<td>4,44</td>
<td>4,49</td>
<td>3,94</td>
</tr>
<tr>
<td>Color Spectrum</td>
<td>3,70</td>
<td>4,14</td>
<td>4,17</td>
<td>4,09</td>
<td>3,56</td>
<td>3,99</td>
<td>3,66</td>
<td>4,15</td>
<td>4,14</td>
</tr>
<tr>
<td>Watermarks</td>
<td>4,33</td>
<td>4,66</td>
<td>3,59</td>
<td>4,43</td>
<td>3,07</td>
<td>4,29</td>
<td>4,79</td>
<td>4,35</td>
<td>4,46</td>
</tr>
<tr>
<td>Security Thread</td>
<td>4,28</td>
<td>4,87</td>
<td>4,60</td>
<td>4,35</td>
<td>3,41</td>
<td>3,72</td>
<td>4,56</td>
<td>4,31</td>
<td>4,31</td>
</tr>
<tr>
<td>Security Fibers, etc.</td>
<td>3,61</td>
<td>4,42</td>
<td>0,36</td>
<td>4,59</td>
<td>3,04</td>
<td>4,37</td>
<td>4,07</td>
<td>4,65</td>
<td>4,65</td>
</tr>
<tr>
<td>Design &amp; Theme</td>
<td>3.08</td>
<td>4.24</td>
<td>4.14</td>
<td>3.98</td>
<td>3.79</td>
<td>3.66</td>
<td>3.66</td>
<td>4.30</td>
<td>4.16</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>Intaglio</td>
<td>4.61</td>
<td>4.71</td>
<td>4.30</td>
<td>4.49</td>
<td>4.10</td>
<td>4.37</td>
<td>4.32</td>
<td>4.69</td>
<td>4.69</td>
</tr>
<tr>
<td>Orloff Printing</td>
<td>4.18</td>
<td>0.18</td>
<td>0.00</td>
<td>4.64</td>
<td>0.00</td>
<td>0.24</td>
<td>1.86</td>
<td>4.68</td>
<td>4.82</td>
</tr>
<tr>
<td>Iris Printing</td>
<td>3.74</td>
<td>4.13</td>
<td>1.41</td>
<td>4.33</td>
<td>4.04</td>
<td>3.21</td>
<td>3.99</td>
<td>4.54</td>
<td>4.58</td>
</tr>
<tr>
<td>Numbering</td>
<td>3.95</td>
<td>3.53</td>
<td>3.35</td>
<td>3.92</td>
<td>4.34</td>
<td>3.89</td>
<td>4.20</td>
<td>4.24</td>
<td>4.31</td>
</tr>
<tr>
<td>Signature</td>
<td>2.43</td>
<td>1.07</td>
<td>1.03</td>
<td>1.32</td>
<td>4.08</td>
<td>3.90</td>
<td>4.03</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Emblem of the Bank</td>
<td>3.40</td>
<td>3.64</td>
<td>3.44</td>
<td>3.81</td>
<td>4.24</td>
<td>3.41</td>
<td>3.60</td>
<td>4.26</td>
<td>4.05</td>
</tr>
<tr>
<td>Vis.-Impaired Marks</td>
<td>3.82</td>
<td>0.86</td>
<td>0.21</td>
<td>0.00</td>
<td>1.73</td>
<td>1.84</td>
<td>3.93</td>
<td>4.30</td>
<td>4.38</td>
</tr>
<tr>
<td>Microprinting</td>
<td>4.74</td>
<td>4.55</td>
<td>4.36</td>
<td>4.76</td>
<td>3.81</td>
<td>4.25</td>
<td>4.44</td>
<td>4.69</td>
<td>4.69</td>
</tr>
<tr>
<td>Micropattern &amp; Trap</td>
<td>4.03</td>
<td>4.38</td>
<td>4.34</td>
<td>4.42</td>
<td>3.97</td>
<td>3.82</td>
<td>3.82</td>
<td>4.20</td>
<td>4.21</td>
</tr>
<tr>
<td>Background Grid</td>
<td>3.86</td>
<td>4.25</td>
<td>4.28</td>
<td>4.34</td>
<td>3.44</td>
<td>3.69</td>
<td>4.09</td>
<td>4.13</td>
<td>4.23</td>
</tr>
<tr>
<td>Graphic Elements</td>
<td>3.06</td>
<td>4.06</td>
<td>4.29</td>
<td>3.93</td>
<td>4.15</td>
<td>3.96</td>
<td>3.99</td>
<td>4.31</td>
<td>4.27</td>
</tr>
<tr>
<td>See-through Register</td>
<td>3.89</td>
<td>4.72</td>
<td>4.36</td>
<td>4.18</td>
<td>0.00</td>
<td>4.22</td>
<td>4.29</td>
<td>2.34</td>
<td>0.00</td>
</tr>
<tr>
<td>Latent Image (Kipp)</td>
<td>4.04</td>
<td>2.82</td>
<td>2.47</td>
<td>4.40</td>
<td>0.00</td>
<td>0.00</td>
<td>4.00</td>
<td>4.01</td>
<td>3.93</td>
</tr>
<tr>
<td>MVC Feature</td>
<td>4.05</td>
<td>0.00</td>
<td>0.00</td>
<td>4.31</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>4.56</td>
<td>4.61</td>
</tr>
<tr>
<td>OVI</td>
<td>4.26</td>
<td>1.66</td>
<td>3.90</td>
<td>0.00</td>
<td>3.34</td>
<td>0.00</td>
<td>4.13</td>
<td>4.04</td>
<td>3.91</td>
</tr>
<tr>
<td>Iridescent Ink</td>
<td>3.93</td>
<td>3.60</td>
<td>3.41</td>
<td>0.00</td>
<td>4.46</td>
<td>3.40</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Metallized Ink</td>
<td>3.08</td>
<td>3.98</td>
<td>0.29</td>
<td>1.61</td>
<td>3.25</td>
<td>0.24</td>
<td>3.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>OVD</td>
<td>4.31</td>
<td>4.69</td>
<td>4.46</td>
<td>4.06</td>
<td>0.00</td>
<td>4.79</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Foil Embossing</td>
<td>3.14</td>
<td>0.29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Blind Embossing</td>
<td>2.97</td>
<td>2.61</td>
<td>0.00</td>
<td>3.99</td>
<td>0.00</td>
<td>0.00</td>
<td>2.38</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>MicroPerf</td>
<td>4.14</td>
<td>0.00</td>
<td>0.00</td>
<td>4.27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>4.74</td>
<td>4.75</td>
</tr>
<tr>
<td>UV-Image</td>
<td>4.59</td>
<td>4.70</td>
<td>3.71</td>
<td>4.46</td>
<td>2.49</td>
<td>4.65</td>
<td>4.83</td>
<td>3.59</td>
<td>3.54</td>
</tr>
<tr>
<td>IR “M” feature</td>
<td>4.11</td>
<td>0.89</td>
<td>1.19</td>
<td>0.36</td>
<td>1.44</td>
<td>0.00</td>
<td>3.68</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Magnetic Image</td>
<td>4.19</td>
<td>3.63</td>
<td>0.91</td>
<td>2.21</td>
<td>3.96</td>
<td>3.41</td>
<td>2.67</td>
<td>2.21</td>
<td>2.25</td>
</tr>
<tr>
<td>Other (see remarks)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Average Mark</strong></td>
<td>3.22</td>
<td>2.64</td>
<td>3.30</td>
<td>2.54</td>
<td>2.82</td>
<td>3.39</td>
<td>3.25</td>
<td>3.16</td>
<td></td>
</tr>
<tr>
<td><strong>General Mark</strong> (WAQI)</td>
<td>3.29</td>
<td>2.75</td>
<td>3.40</td>
<td>2.53</td>
<td>2.89</td>
<td>3.43</td>
<td>3.38</td>
<td>3.28</td>
<td></td>
</tr>
<tr>
<td><strong>Average Rating</strong></td>
<td>1.71</td>
<td>3.00</td>
<td>1.29</td>
<td>5.00</td>
<td>3.93</td>
<td>1.93</td>
<td>1.57</td>
<td>2.57</td>
<td></td>
</tr>
</tbody>
</table>

These results represent sufficient interest, first of all, from point of view, that during that case study the offered technique has been checked up and fulfilled in practice, a lot of details arising during research were found out.

Due to kind help of experts (actually, the research of banknotes and filling the inquiry map takes more than 3 hours) the opportunity has appeared not only to compare various banknotes (which task, as a matter of fact, was not the primary goal), but, the most important, to evaluate importance and quality of realization for various security features - from the point of view of each expert group and as a whole. These data are of interest first of all for manufacturers of banknotes, however, as well as for Central and National banks.

In this speech the problem of detailed analysis of received results isn’t put forward, we mention only some of those. First of all, it is evident, that according to experts’ assessment on the
average, the most qualitative are the sample “Leo Tolstoy” of Goznak and the banknote of 200 Slovak krones.

It might be interesting to look at quality estimation average WAQI for different Expert Groups (Table 2).

<table>
<thead>
<tr>
<th>Expert Group</th>
<th>G&amp;D BN200</th>
<th>DLRG LeMond</th>
<th>Goznak Tolstoy</th>
<th>USD 20</th>
<th>EUR 10</th>
<th>SKK 200</th>
<th>RUR 500</th>
<th>RUR 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial Banks</td>
<td>3,348</td>
<td>2,843</td>
<td>3,427</td>
<td>2,537</td>
<td>2,893</td>
<td>3,423</td>
<td>3,353</td>
<td>3,250</td>
</tr>
<tr>
<td>Bank of Russia</td>
<td>3,302</td>
<td>2,819</td>
<td>3,655</td>
<td>2,687</td>
<td>3,133</td>
<td>3,790</td>
<td>3,517</td>
<td>3,432</td>
</tr>
<tr>
<td>Forensic Science Center</td>
<td>3,117</td>
<td>2,429</td>
<td>3,064</td>
<td>2,355</td>
<td>2,636</td>
<td>3,071</td>
<td>3,310</td>
<td>3,211</td>
</tr>
</tbody>
</table>

We can apply various complex cost indexes to derive financial and economic assessment of given series of banknote quality, elementary of which are: Relative Cost of Banknote Production \( C_{rchp} \) and Specific Banknote Quality Index \( Q_{sbq} \). Nobody tells us production costs. But we can calculate, at least approximately, Specific Banknote Quality Index of estimated banknotes, having taken for convenience a nominal value of note in rubles with multiplication of result to 100.

If we want to group quality performances of a banknote on the base of main production phases and kinds of protection we can see the biggest “weight” of physical-chemical protection (otherwise different inks usage). Although all parameters are approximately identical on the average actually. Thus, experts “consider” that from the point of view of banknote quality it is necessary to improve all three options (banknote paper, printing methods and inks).

Another point we could mention as well are the features of banknote quality, determined as most important by different expert groups by rank. As we see, the attitude to definition of the most important banknote quality features (as well as the less important ones) in different expert groups quite differ.

After some very interesting discussions at “Banknote 2008” Conference in Washington I made some additional calculations. The main idea of those was to estimate importance of different security features from point of view to check those by most popular testing methods (manually, visually, by lens, in UV and IR spectrum, by magnetic detectors). The most important were security thread, numeration, Orloff printing and microperforation. The less important – metalized ink, foil embossing and Emblem of the Bank design.

The results of assessment of different testing methods importance from point of view their usage in checking of security features are shown at Fig. 2 and 3.

![Fig.2. Assessment of Testing Methods Importance](image-url)
Last year this approach was applied again due to introduction of number of newest security features in RUR 1000 and new series of National Bank of Kyrgyz Republic (NBKR) banknotes. Third round of assessment has been undertaken in January and February 2011 where 25 specialists evaluate 38 parameters of 4 circulating banknotes only. As before, experts represented three target groups distinguished as under the approach to assessment, working experience and level of technical equipment: forensic experts, Central Bank experts and cashiers from different commercial banks. A new and interesting difference was participation of NBKR experts.

The results of all 3 rounds of banknote quality assessment provide us good statistics (3720, 4284 and 4750 assessments correspondingly) and possibility to determine effectiveness of introduction of new design and different new security features from the point of view of different expert target groups and in a whole.

After invitation to speak here at SMTDA 2012 International Conference I continue to improve “the tool”. One of problems was serious difference between assessments in Cashiers Target Group. So I calculate Importance Mark by several statistical functions (Table 3).

Table 3

<table>
<thead>
<tr>
<th>Importance Mark for Commercial Banks</th>
<th>Statistical Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode</td>
<td>Median</td>
</tr>
<tr>
<td>Feel of paper</td>
<td>5,00</td>
</tr>
<tr>
<td>Size of banknote</td>
<td>4,00</td>
</tr>
<tr>
<td>Colour spectrum</td>
<td>5,00</td>
</tr>
<tr>
<td>Watermarks, including their filigree</td>
<td>5,00</td>
</tr>
<tr>
<td>Security threads (design, symbols, luminescence, etc.)</td>
<td>5,00</td>
</tr>
<tr>
<td>Security fibers, confetti, etc. (design, luminescence)</td>
<td>4,00</td>
</tr>
<tr>
<td>Design and layout (clarity, saturation)</td>
<td>5,00</td>
</tr>
<tr>
<td>Intaglio relief</td>
<td>5,00</td>
</tr>
<tr>
<td>Orloff printing</td>
<td>4,00</td>
</tr>
<tr>
<td>Iris printing (incl. 2D-Iris)</td>
<td>4,00</td>
</tr>
<tr>
<td>Numbering (system, type size, positions)</td>
<td>5,00</td>
</tr>
<tr>
<td>Signature</td>
<td>2,00</td>
</tr>
<tr>
<td>Emblem, coat of arms etc.</td>
<td>1,00</td>
</tr>
<tr>
<td>Marks for visually impaired</td>
<td>4,00</td>
</tr>
<tr>
<td>Microprinting</td>
<td>5,00</td>
</tr>
<tr>
<td>Special raster patterns</td>
<td>4,00</td>
</tr>
</tbody>
</table>
This table provides us information where “Tastes are Differ” and where those are the same. And minimize that influence to final result.

That next step is good basis to move forward. Especially because new security features appears almost every month: overt and covert, different combinations of design and protection. And we see the challenges for global currencies in nearest future. So new approaches and new case studies are needed.

### 5. Conclusions

To sum it up I would like to say, that represented approach assumes diverse variants of formation both Expert Group(s) and Workgroup; more or less detailed and deep development of parameters’ system for evaluation. All of those anyway will depend on the customer – Central (National) Bank of the country or whoever. At the same time, we need to keep in mind the Indeterminacy Principle of Heisenberg - in sense that the more details we try to put into the model, the less authentic can be appeared results. Or, as the proverb says: «The Best is the enemy of the Good».

Unconditionally, the described verification of approach is largely experimental as it was conducted on a free basis, and mainly due to serious attitude and respect of the experts to their trade. The time to don't distract them from fulfillment of current duties, the time for an explanation of estimation technique and features of samples was not always sufficed. At the same time, the experts have told after all, that this work was interesting and useful for them as well.

The considered approach can be applied both to the analysis of existing banknote quality, as well as at the development stage of new series of banknotes for the multicriteria analysis and optimization of parameters. Moreover it can become a quite good tool for the analysis of

| Micropatterns and graphic traps | 4,00 | 4,00 | 4,06 | 3,96 | 0,66 | 0,54 |
| OMROM Rings (EURion Constellation) | 4,00 | 4,00 | 3,42 | 3,01 | 2,15 | 1,24 |
| Background grids | 5,00 | 4,00 | 3,90 | 3,69 | 1,07 | 0,73 |
| Graphic elements of a decor | 3,00 | 4,00 | 3,95 | 3,89 | 0,51 | 0,59 |
| See-through register | 3,00 | 4,00 | 3,99 | 3,92 | 0,59 | 0,62 |
| Latent images (Kippeffekte) | 4,00 | 4,00 | 3,79 | 3,70 | 0,72 | 0,66 |
| MVC (MVC +) effect | 4,00 | 4,00 | 4,16 | 4,10 | 0,48 | 0,53 |
| OVI, OVMI | 4,00 | 4,20 | 4,29 | 4,26 | 0,30 | 0,43 |
| Iridescent inks | 5,00 | 4,30 | 3,81 | 3,52 | 1,59 | 1,01 |
| Metallized inks | 4,00 | 3,10 | 3,24 | 3,03 | 1,11 | 0,86 |
| Kinegrams and holograms (OVD) (stripe or patch) | 5,00 | 4,80 | 4,38 | 4,27 | 0,80 | 0,69 |
| Foil embossing | 1,00 | 3,00 | 2,44 | 2,11 | 1,46 | 1,06 |
| Blind embossing | 3,00 | 3,20 | 3,40 | 3,33 | 0,46 | 0,55 |
| Microperforation | 4,00 | 4,00 | 4,14 | 3,99 | 0,98 | 0,71 |
| UV–images (including UV-bi-fluorescence) | 5,00 | 4,50 | 4,32 | 4,26 | 0,48 | 0,58 |
| IR–images (easiness to check) | 5,00 | 5,00 | 4,93 | 4,92 | 0,03 | 0,12 |
| Anti-Stokes effect | 1,00 | 1,50 | 2,06 | 1,74 | 1,52 | 1,09 |
| IR special element “M” (“flickering, blinking IR”) | 4,80 | 4,00 | 3,36 | 2,89 | 2,48 | 1,37 |
| Magnetic image | 3,00 | 3,00 | 3,51 | 3,40 | 0,83 | 0,76 |

SMTDA 323
banknote production competitiveness in different regions worldwide (if we don't take into account some political factors).

In turn, Central or National Bank which is choosing manufacturer of banknotes for the country can use the similar approach as the additional tool.

For development of international standards in the field of banknote production and protection it would be possible to create the Commission of experts from representatives of leading banknote manufacturers, Central and National banks experts and experts of the Interpol under aegis of one of international organizations (Intergraf, for example) or maybe some International Conferences.

To finish my speech I’d like to say proper words from “Currency News” monthly (January 2009 Issue) Editorial: “Currency needs the best of protection at the best of times, and even better protection when times are hard”.

Many thanks for your attention.
Differentiation of Mortality Depending upon Type of the Settlement

Alla Ivanova¹, Elena Zemlyanova², Olga Antonova³, Svetlana Nikitina⁴

¹ Federal Public Health Institute, Moscow, Russia
e-mail: ivanova-home@yandex.ru

² Federal Public Health Institute, Moscow, Russia
e-mail: zemlianova@mednet.ru

³ Federal State Statistics Service, Moscow, Russia
e-mail: antonova_o@gks.ru

⁴ Federal State Statistics Service, Moscow, Russia
e-mail: nikitina_s@gks.ru

Abstract. Urban population in Russia is heterogeneous in its mortality levels. Mortality is minimal in administrative centers of Federal subjects and maximum in little towns with population number under 10,000 inhabitants. Though, no visible differences in mortality were found between urban settlements with population number 10,000 – 100,000 inhabitants. This evidences about rapprochement of quality of life in mid-sized and little towns. Such rapprochement is not due to growth of quality of life in little towns but its decline in mid-sized towns. Little towns with population number under 10,000 inhabitants and urban-type settlements become an area of higher risk in all population groups: from children to elderly. Population of big cities and administrative centers in all age groups is on the “winning side”. Though, situation in children and youth is better in big cities, but for ages over 30 years situation is better in administrative centers. It appears that the status of capital even regional brings not only advantages in quality of life but additional risks for adolescent and young ages. Such additional risks are minimal or less pronounced in provincial cities even big ones. Depending upon type of urban settlement not only level of mortality is different but its cause structure as well. Vector of mortality decline from little towns to big cities was shown for all main causes of death except neoplasms. This raises a problem of quality of mortality causes diagnostics especially in little towns.

Introduction

In Russia existing essential regional variation of mortality amounts 15 years of male life expectancy and 10 years of female life expectancy. Though inter-regional mortality differences sometimes are the same essential. Besides they are characteristic both for favourable territories with low mortality and for unfavourable territories as well. So, in Khanty-Mansiisk autonomous district – Yugra being one of the ten Russian regions with maximum life expectancy (65.7 years – males, 75.7 years – females in 2010) differences between cities/towns and districts amounted 14.4 years in males and 9.5 years in females. In Smolensk oblast being one of the ten Russian regions with minimal life expectancy (59.5 years – males, 73.3 years – females in 2010)
the corresponding differences amounted 13.0 years in males and 10.9 years in females. Such essential differences are determined by the mortality differences between rural districts and regional center. Intermediate position is taken by the districts including towns varying by population number. The aim of this study is to specify mortality differences in urban population depending upon type of settlement and population number.

Data and methods

The study was conducted using official mortality data provided by Federal State Statistics Service. Mortality data were differentiated by types of urban settlements: urban-type settlements, little towns under 10,000 inhabitants, little towns with 10,000-19,999 inhabitants, mid-sized towns with 20,000-49,999 inhabitants, mid-sized towns with 50,000-99,999 inhabitants and cities with population number over 100,000 inhabitants. Separately, not depending upon population number we analyzed administrative centers of Federal subjects. Thus, we selected 7 groups of urban settlements with consideration of their type and population number.

The study was conducted using data for the years 2008-2009. Accordingly, data about number of deaths for those two years served as a numerator, and a duplicated number of population for 01.01.2009 for corresponding groups of urban settlements served as a denominator.

So, the object of the study was formed of territories of the Central Federal District where about one quarter of Russian population lives. We excluded from our study Moscow and Moscow oblast. Moscow was excluded because of its specific status of megalopolis as a subject of Federation, and Moscow oblast was excluded because it hasn’t its own administrative center. As a result, we analyzed data on 16 subjects of Federation and covered about 10 million of Russian population.

Results

At an average, population mortality for all territories of the Central Federal District occurred minimal in administrative centers of the subjects of Federation: 1739.3 per 100,000 males and 832.1 per 100,000 females. Maximum observations were characteristic for little towns (under 10,000 inhabitants) – 2095.5 per 100,000 males and 977.6 per 100,000 females (table I). Thus, we confirmed expected variation vector. Variance amounts about 20% in males and 17.5% in females. For the first glance, it isn’t too much. But if we compare the best and the worst mortality levels with regional distribution of Russian mortality it occurs that in males indices of administrative centers the most closely correspond to mortality level in the regions taking 18th rank in males and 19th rank in females. At the same time, indices for little towns (under 10,000 inhabitants) the most closely
correspond to mortality level in the regions taking 72\textsuperscript{th} rank in males and 77\textsuperscript{th} rank in females.

Table I. Population mortality depending upon type of the settlement and number of population in urban settlements in the Central Federal District of the Russian Federation, years 2008-2009, standardized rates, European standard, per 100,000

<table>
<thead>
<tr>
<th>Type of Settlement</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban-type settlements</td>
<td>1960.2</td>
<td>920.8</td>
</tr>
<tr>
<td>Little towns:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>under 10,000</td>
<td>2095.5</td>
<td>977.6</td>
</tr>
<tr>
<td>10,000-19,999</td>
<td>1999.8</td>
<td>949.0</td>
</tr>
<tr>
<td>Mid-sized towns:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20,000-49,999</td>
<td>1984.2</td>
<td>921.8</td>
</tr>
<tr>
<td>50,000-99,999</td>
<td>1993.8</td>
<td>920.5</td>
</tr>
<tr>
<td>Big cities:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100,000 and over</td>
<td>1820.1</td>
<td>872.3</td>
</tr>
<tr>
<td>Center of the subject of Federation</td>
<td>1739.3</td>
<td>832.1</td>
</tr>
</tbody>
</table>

It could be expected that the highest levels of mortality would be found in urban-type settlements because their urban status is reasonably formal. As a rule, such settlements arose around big industrial enterprises, and level of social infrastructure development there was traditionally behind the same in towns and cities. At the same time, population mortality in urban-type settlements occurred to be close to the levels of mid-sized towns (20,000-49,999 inhabitants). Probably it could be associated with gradual social and economic degradation of towns, negative net migration, desertion connected with lack of job vacancies and absence of own sources for development. The second peculiarity is that no visible mortality differences were found between urban settlements with population from 10,000 to 100,000 inhabitants. In little towns (10,000-19,999 inhabitants) mortality levels were 1999.8 per 100,000 males and 949.0 per 100,000 females; and in mid-sized towns (50,000-99,999 inhabitants) mortality indices were 1993.8-920.5 per 100,000 correspondingly. Such results evidence about commonness of factors determining high mortality in those types of urban settlements. In other words, quality of life in mid-sized and little towns moves closer, and it isn’t improving in little towns but it is getting worse in mid-sized towns. Only big cities give advantages in mortality decline to their inhabitants providing broader opportunities for education and work, receiving of social support and medical care. Inhabitants of administrative centers of subjects of Federation are even in better situation. However their gain in mortality in comparison with the inhabitants of big cities is less than advantage of the inhabitants of big cities in comparison with the inhabitants of mid-sized and little towns. Thus, among essential factors forming conditions for mortality decline there is not only the status of administrative center but, first of all, population number in town or city determining economic and social potential of its inhabitants.
Not only level of mortality but its cause structure differs depending upon type of urban settlement (fig. 1).

![Death causes structure depending upon type and population number of the urban settlement in the Central Federal District, years 2008-2009, standardized rates, European standard, per 100,000](image)

Fig. 1. Death causes structure depending upon type and population number of the urban settlement in the Central Federal District, years 2008-2009, standardized rates, European standard, per 100,000

Variation of mortality levels is determined by cardio-vascular diseases whose input into mortality structure is more than 50% in males and about 65% in females. Higher mortality from this class of causes is registered in little towns and on the contrary, lower levels are registered in big cities and centers of subjects of Federation. Similar regularities although not so distinct are manifested in the case of respiratory diseases, digestive diseases, infections
and accidents and traumas. The only important cause of death showing inverse relation is malignant neoplasms when lower mortality levels are characteristic for population of little towns (both sexes) and urban-type settlements (females), and the highest mortality levels are characteristic for females living in administrative centers and for males living in whole entity of urban settlements with population number exceeding 50,000 inhabitants. Neoplasms falling out from mortality vector of urban population raise the question of completeness and reliability of death causes coding even in urban settlements. More strict criteria for establishment of neoplasm as a death cause maybe are leading to situation when a certain number of deaths from neoplasms are being masked by other diagnosis. First of all, it is possible when there are no capacities for qualified post mortem examination. So, quality of diagnostics, especially in little towns is questionable. Let’s see what age groups form mortality vector of urban population (tab. 2).

### Table II. Age-specific death rates depending upon type and population number of the urban settlement in The Central Federal District of the Russia Federation, years 2008-2009, per 100,000

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Males</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Females</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-14</td>
<td>15-29</td>
<td>30-44</td>
<td>45-59</td>
<td>≥60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban-type settlements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>106.8</td>
<td>365.9</td>
<td>893.8</td>
<td>2265.8</td>
<td>7955.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Little towns:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>under 10,000</td>
<td>145.9</td>
<td>271.6</td>
<td>1001.0</td>
<td>2502.3</td>
<td>8349.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10,000-19,999</td>
<td>99.8</td>
<td>331.9</td>
<td>891.4</td>
<td>2326.0</td>
<td>8152.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid-sized towns:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20,000-49,999</td>
<td>104.5</td>
<td>310.5</td>
<td>896.1</td>
<td>2252.9</td>
<td>8054.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50,000-99,999</td>
<td>117.3</td>
<td>332.9</td>
<td>862.2</td>
<td>2241.0</td>
<td>8252.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Big cities:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100,000 and over</td>
<td>79.9</td>
<td>230.0</td>
<td>776.8</td>
<td>2043.1</td>
<td>7431.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Center of the subject of Federation</td>
<td>93.2</td>
<td>243.3</td>
<td>752.1</td>
<td>1932.7</td>
<td>7169.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban-type settlements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85.8</td>
<td>102.9</td>
<td>273.5</td>
<td>690.9</td>
<td>5321.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Little towns:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>under 10,000</td>
<td>53.0</td>
<td>135.8</td>
<td>264.1</td>
<td>807.0</td>
<td>5625.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10,000-19,999</td>
<td>72.4</td>
<td>99.3</td>
<td>291.3</td>
<td>717.1</td>
<td>5471.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid-sized towns:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20,000-49,999</td>
<td>77.9</td>
<td>91.4</td>
<td>272.7</td>
<td>661.2</td>
<td>5299.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50,000-99,999</td>
<td>69.4</td>
<td>96.3</td>
<td>250.4</td>
<td>678.1</td>
<td>5335.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Big cities:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100,000 and over</td>
<td>55.3</td>
<td>54.1</td>
<td>248.8</td>
<td>625.1</td>
<td>4902.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Center of the subject of Federation</td>
<td>73.6</td>
<td>67.5</td>
<td>224.7</td>
<td>577.3</td>
<td>4735.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For all age groups little towns (under 10,000 inhabitants) and urban-type settlements become the areas of higher risk. On the contrary minimal levels of mortality are registered in big cities and administrative centers of subjects of Federation. Notably, situation for children and youth is better in big cities, and for ages over 30 years better situation is in administrative centers. It appears that the status of capital even regional brings not only advantages in quality of life but additional risks for adolescent and young ages. Such additional risks are minimal or less pronounced in provincial cities even big ones. In the case of older age groups, opportunities for self-realization (in working ages) and higher availability of qualified medical care (in elderly) determine mortality gain for inhabitants of administrative centers.

Conclusion

Complicated structure of urban settlements existing in Russia determines essential heterogeneity of urban population mortality. It is shown by an example of the Central Federal District of the Russian Federation that mortality of population in little towns (under 10,000 inhabitants) is higher than in administrative centers by 20% in males and by 17.5% in females. Such variance is determined by all main death causes except neoplasms and all population groups – from children to elderly. Variation of mortality in every region is determined by a proportion of inhabitants of little towns and big cities while differences between little and mid-sized towns (from 10,000 to 99,999 inhabitants) are of little importance. This evidences about commonness of factors determining high levels of mortality in such types of settlements. In other words, rapprochement of quality of life in mid-sized and little towns is not due to its improvement in little towns but its decline in mid-sized towns.
Stochastic data envelopment analysis network models

Josef Jablonsky
Department of Econometrics, Faculty of Informatics and Statistics,
University of Economics, Prague, Czech Republic
Email: jablon@vse.cz

Abstract: Data envelopment analysis (DEA) is a non-parametric technique for evaluation of relative efficiency of decision making units described by multiple inputs and outputs. It is based on solving linear programming problems. Since 1978 when basic DEA model was introduced many its modifications were formulated. Among them are two or multi-stage models with series or parallel structure (network models). Standard DEA models are based on deterministic inputs and outputs. The paper deals with DEA network models under the assumption that the inputs and/or outputs are continuous random variables. Under this assumption the efficiency scores of production units are random variables as well. Several approaches for description of random efficiency scores were developed for standard DEA models but only few for models with network structure. They are mostly based on formulation of linear optimization problems. Another methodological approach for stochastic DEA models is simulation. The paper compares results given by simulation experiments with several definitions of linear programming models.

Keywords: Data envelopment analysis, network models, DEA models with imprecise data, efficiency

1 Introduction

Data envelopment analysis (DEA) is a non-parametric technique for evaluation of relative efficiency of decision making units described by multiple inputs and outputs. Let us suppose that the set of decision making units (DMUs) contains \( n \) elements. The DMUs are evaluated by \( m \) inputs and \( r \) outputs with inputs and outputs values \( x_{ij}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n \) and \( y_{kj}, \quad k = 1, 2, \ldots, r, \quad j = 1, 2, \ldots, n \), respectively. The efficiency of the \( q \)-th DMU can be expressed as the weighted sum of outputs divided by the weighted sum of outputs with weights reflecting the importance of single inputs/outputs \( v_i, \quad i = 1, 2, \ldots, m \) and \( u_k, \quad k = 1, 2, \ldots, r \) as follows:

\[
\theta_q = \frac{\sum_{k=1}^{r} u_k y_{kq}}{\sum_{i=1}^{m} v_i x_{iq}}.
\]

(1)

Standard CCR input oriented DEA model formulated by Charnes et al. (1978) consists in maximization of efficiency score (1) of the DMU \( q \) subject to constraints that efficiency scores of all other DMUs are lower or equal than 1. The linearized form of this model is as follows:
Maximize \( \theta_q = \sum_{k=1}^{r} u_k y_{kq} \)

subject to

\[
\begin{align*}
\sum_{i=1}^{m} v_i x_{iq} &= 1, \\
\sum_{k=1}^{r} u_k y_{kj} - \sum_{i=1}^{m} v_i x_{ij} &\leq 0, \quad j = 1, 2, \ldots, n, \\
u_k, v_i &\geq \varepsilon, \quad k = 1, 2, \ldots, r, i = 1, 2, \ldots, m.
\end{align*}
\]

If the optimal value of the model (2) \( \theta_q^* = 1 \) then the DMU \( q \) is CCR efficient and it is lying on the CCR efficient frontier, otherwise the unit is not CCR efficient. The model (2) is often referenced as primal CCR model. Its dual form is sometimes more convenient and its mathematical model is as follows:

Minimize \( \theta_q \)

subject to

\[
\begin{align*}
\sum_{j=1}^{n} x_{ij} \lambda_j + s^-_i &= \theta_q x_{iq}, \quad i = 1, 2, \ldots, m, \\
\sum_{j=1}^{n} y_{kj} \lambda_j - s^+_k &= y_{kq}, \quad k = 1, 2, \ldots, r, \\
\lambda_j &\geq 0, \quad j = 1, 2, \ldots, n,
\end{align*}
\]

where \( \lambda_j, i = 1, 2, \ldots, n \) are weights of DMUs, \( s^-_i, i = 1, 2, \ldots, m \), and \( s^+_k, k = 1, 2, \ldots, r \) are slack (surplus) variables and \( \theta_q \) is the efficiency score of the DMU \( q \) which expresses necessary reduction of inputs in order this unit becomes efficient. The models (2) and (3) are CCR models with input orientation, i.e. they look for reduction of inputs in order to reach the efficient frontier. The output oriented modification of the presented models is straightforward. The BCC models under variable returns to scale assumptions originally presented by Banker et al. (1984) extend the formulation (3) by convexity constraint \( \sum \lambda_j = 1 \).

The presented basic DEA models measure efficiency of a transformation of \( m \) inputs into \( r \) outputs in one stage and under an assumption that all data are deterministic but the production process is often more complex and the data may be stochastic. The paper formulates two-stage network model with serial structure where outputs of the first stage are inputs of the second stage and offers its solution under assumption that all inputs and outputs are random variables.

The paper is organized as follows. The next section contains basic formulation of DEA network model with random data and discusses possibility of its solution using optimization and simulation models. Section 3 presents results of numerical experiments based on a real data set – evaluation of efficiency of bank branches.
Final part of the paper summarizes presented results and discusses directions for future research.

2 DEA network models with random data

The models (2) and (3) measure the relative efficiency of one-stage transformation of \(m\) inputs into \(r\) outputs. The transformation of inputs into final outputs can be taken as a two- or several-stage process. The inputs of the first stage are transformed into its outputs and all or at least some of these outputs are utilized as inputs of the second stage that are using for production of final outputs. Let us denote the input values of the first stage \(x_{ij}, i = 1,2,\ldots,m, j = 1,2,\ldots,n\) and the output values of the first stage \(y_{ij}, i = 1,2,\ldots,r, j = 1,2,\ldots,n\). Supposing that all outputs of the first stage are taken as inputs of the second stage and that the final output values are \(z_{ij}, i = 1,2,\ldots,p, j = 1,2,\ldots,n\). Two-stage DEA models are widely analyzed and discussed within professional community. Theoretical issues can be found e.g. in (Liang et al., 2011). Among numerous case studies can be mentioned papers (Jablonsky, 2009) and (Paradi et al., 2011). Two-stage DEA model under constant returns to scale assumption can be formulated according to Chen et al. (2009) as follows:

Minimize

\[
\theta_q - \phi_q
\]

subject to

\[
\sum_{j=1}^{n} x_{ij} \lambda_j \leq \theta_q x_{iq}, \quad i = 1,2,\ldots,m,
\]

\[
\sum_{j=1}^{n} y_{kj} \lambda_j \geq \tilde{y}_{kq}, \quad k = 1,2,\ldots,r,
\]

\[
\sum_{j=1}^{n} y_{kj} \mu_j \leq \tilde{y}_{kq}, \quad k = 1,2,\ldots,r,
\]

\[
\sum_{j=1}^{n} z_{lj} \mu_j \geq \phi_q \tilde{z}_{lq}, \quad l = 1,2,\ldots,p,
\]

\[
\theta_q \leq 1, \phi_q \geq 1,
\]

\[
\lambda_j \geq 0, \mu_j \geq 0, \quad j = 1,2,\ldots,n,
\]

where \(\lambda_j\) and \(\mu_j, j = 1,2,\ldots,n\), are weights of the DMUs in the first and second stage, \(\theta_q\) and \(\phi_q\) efficiency scores of the DMU \(q\) in the first and second stage and \(\tilde{y}_{kq}\) are variables to be determined. The DMU \(q\) is recognized as efficient by model (4) if the efficiency scores in both stages are \(\theta_q = 1\) and \(\phi_q = 1\) respectively, and the optimal objective value of the presented model is 0. The inefficient units can be ranked relatively by the following geometric average efficiency measure:

\[
e_q = (\theta_q / \phi_q)^{1/2}
\]
Inputs and outputs in both stages usually reflect past values of the DMUs. That is why the model (4) supposes that the inputs and outputs of the units are given as deterministic values. For evaluation and estimation of future efficiency of the DMUs it can be useful to consider inputs and outputs as random variables. They can be given as interval values or more generally as random variables with defined continuous probabilistic distribution. Approaches for dealing with random data in DEA can be divided into two groups – optimization and simulation methods. Optimization approaches are based on solving one or several linear programs and result to an index or indices for each DMU that can be used for their ranking. One of the optimization approaches is presented in Despotis and Smirlis (2002). It supposes that the corresponding values for inputs and outputs are continuous variables with uniform distribution defined over intervals $x_{iq} \in \langle x_{Liq}, x_{Uiq} \rangle$, $y_{iq} \in \langle y_{Liq}, y_{Uiq} \rangle$ and $z_{iq} \in \langle z_{Liq}, z_{Uiq} \rangle$. Efficiency scores of the DMUs under the assumption of interval inputs and outputs are random variables defined over an interval. The lower bound for efficiency score in both stages is given by using of the worse inputs and outputs for the evaluated unit DMU $q$ and the best characteristics for all the other units and similarly the upper bound is defined by using the best characteristics of the evaluated unit and the worse ones for all the other units. The optimization model for deriving the lower bound for efficiency score of the unit DMU $q$ in the first stage is as follows:

\begin{align*}
\text{Minimize} & \quad \theta_q \\
\text{subject to} & \quad \sum_{j=1, j \neq q}^{n} (\lambda_{ij} x_{ij}^L + \lambda_{iq} x_{iq}^U + s_{iq}^-) = \theta_{kjq}, \quad i = 1, 2, \ldots, m, \\
& \quad \sum_{j=1, j \neq q}^{n} (\lambda_{ij} y_{ij}^U - \lambda_{iq} y_{iq}^L - s_{iq}^+) = y_{kjq}, \quad k = 1, 2, \ldots, r, \\
& \quad \lambda_{ij} \geq 0, s_{iq}^- \geq 0, s_{iq}^+ \geq 0.
\end{align*}

According to the values of the lower and upper bounds of efficiency scores in each of two stages the DMUs can be divided into three subsets $E^1$, $E^2$ and $E^3$:

- $E^1$: DMUs always efficient – this subset contains units that are efficient in any case, i.e. even their inputs and outputs are on their worst values and the inputs and outputs of other units are on their best bounds.
- $E^2$: DMUs conditionally efficient by suitable adjusting of inputs and outputs of all the units (upper bound of their efficiency score is 1).
- $E^3$: DMUs never efficient (upper bound of their efficiency score is lower than 1).

This approach can lead to quite different results, e.g. a DMU can belong to the set $E^1$ (always efficient) in the first stage and to the set $E^2$ or even $E^3$ (never efficient) in the second stage. In order to evaluate the efficiency of both stages simultaneously using model (4) the optimization model can consider random variables of first stage inputs’ and final outputs and intermediate characteristics use in their average level. The model (4) is then modified according to the model (6), i.e. using lower/upper bounds for characteristics in the first and last group of
constraints. The results of the modified model (4) allow dividing of the units into three classes as above: always efficient (efficient in both stages), never efficient (inefficient in both stages), and conditionally efficient.

One of the disadvantages of the optimisation approach to dealing with random inputs and outputs in DEA models consists in the necessity to consider just interval values (uniform distribution). Except optimisation models simple simulation tools can be used to analyse the presented problem under a more general assumptions. Simulation approach is more time consuming than the optimisation one but it gives much more information that can be useful for a detailed analysis of the problem. This approach can be simply described by the following steps:

1. Generation of all random variables of the model. This step can be simply realised within MS Excel environment by means of built-in functions or VBA procedures.
2. Modified two-stage DEA models (4) are solved with the values generated in the previous step. In our experiments we used LP solver included in the LINGO modelling system which is powerful enough and allows a simple linking with MS Excel sheets.
3. Information from the random trials are processed and evaluated by means of a suitable software tool which can be MS Excel or one of its add-ins (Crystal Ball, @RISK).

Simulation trials give much more information about distribution of efficiency scores of particular units comparing to above described optimization procedure. Results given by both - optimization and simulation – procedure are compared on a simple example in the next section.

3 Computational experiments

Applications of DEA models are numerous. Results of the above formulated models will be illustrated on an example of 67 selling branches of one of the Czech mobile phone operators. The following inputs, intermediate characteristics (outputs of the first stage and inputs of the second one) and final outputs are taken into account.

Inputs:
- Operational expenses (rental costs, wages and overheads), and
- Number of business hours per year is an important characteristic influencing total number of transactions (one of the outputs of the first stage).

Intermediate characteristics:
- Number of transactions of current customers, and
- Number of transactions of new customers.

Outputs:
- Financial contribution of the branch in CZK (Czech crowns), and
- ICCA score measuring the customer satisfaction.

The data for all 67 branches are available with a certain level of uncertainty. That is why the fixed data from 2010 and used for numerical experiments together with
modified set of data. In this modification we suppose that the data are independent continuous random variables with uniform distribution over interval $(0.95x, 1.10x)$, where $x$ is original fixed value. Computational experiments are divided into two groups.

1. Evaluation of efficiency of the branches in two separated stages using model (5) and its modifications under variable returns to scale assumption. Lower and upper bounds for efficiency scores of the branches in both stages are basic results of applied models. The branches can be ranked according to several criteria in this case – e.g. geometric average of maximum or minimum efficiencies in both stages. The results of this approach for a selection of 10 branches are presented in Table 1.

Table 1- Results of efficiency evaluation (two single stages)

<table>
<thead>
<tr>
<th>DMU</th>
<th>1st stage lower</th>
<th>1st stage upper</th>
<th>2nd stage lower</th>
<th>2nd stage upper</th>
<th>Geometric average</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9014</td>
<td>1.0000</td>
<td>0.3382</td>
<td>0.6128</td>
<td>0.6556</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.7139</td>
<td>0.9889</td>
<td>0.6986</td>
<td>1.0000</td>
<td>0.7847</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>0.8948</td>
<td>1.0000</td>
<td>0.6082</td>
<td>1.0000</td>
<td>0.8496</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>0.8572</td>
<td>1.0000</td>
<td>0.5228</td>
<td>0.9757</td>
<td>0.8038</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>0.6191</td>
<td>1.0000</td>
<td>0.4905</td>
<td>0.8589</td>
<td>0.6650</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>0.5876</td>
<td>1.0000</td>
<td>0.5133</td>
<td>0.9546</td>
<td>0.6801</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>0.5662</td>
<td>0.8217</td>
<td>0.9951</td>
<td>1.0000</td>
<td>0.6931</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>0.7996</td>
<td>1.0000</td>
<td>0.3609</td>
<td>1.0000</td>
<td>0.7422</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>0.5745</td>
<td>0.9452</td>
<td>0.4094</td>
<td>0.7141</td>
<td>0.5695</td>
<td>10</td>
</tr>
</tbody>
</table>

According to the results of the first stage there is only one unit (DMU2) that is always efficient, 6 units are conditionally efficient and the remaining ones are always inefficient. Similar conclusions hold for the second stage – again the unit DMU2 is always efficient, other 4 units are conditionally efficient and 5 units are inefficient even they work on their best bound. The pre-last column of Table 1 contains geometric average of two values – simple average of lower and upper values in each stage. The last column presents ranking of the selected units according to the geometric average in the pre-last column. It is really questionable what criterion to use for ranking of units when both stages are separated. It is clear that the most efficient unit is the unit DMU2 that is always efficient in both stages. Next places in ranking are occupied by units that are conditionally efficient at least in one of two stages.

2. Evaluation of efficiency using modified model (4), i.e. considering both stages simultaneously. Optimization approach leads to lower and upper bounds for efficiency scores as above. Simulation approach was realized with uniformly generated data of all units and optimization run with model (4). After 50 trials some information about distribution of efficiency scores in both stages and overall efficiency are given.
Table 2- Results of efficiency evaluation (two-stage model)

<table>
<thead>
<tr>
<th>DMU</th>
<th>optimization lower</th>
<th>optimization upper</th>
<th>simulation (50 trials) min</th>
<th>simulation (50 trials) max</th>
<th>simulation (50 trials) avg</th>
<th>Rank opt</th>
<th>Rank sim</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.6138</td>
<td>0.7233</td>
<td>0.6174</td>
<td>0.7212</td>
<td>0.6635</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.6491</td>
<td>0.8865</td>
<td>0.6814</td>
<td>0.8302</td>
<td>0.7507</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>0.7873</td>
<td>0.9537</td>
<td>0.8067</td>
<td>0.9487</td>
<td>0.8629</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>0.6326</td>
<td>0.8982</td>
<td>0.6908</td>
<td>0.8906</td>
<td>0.7813</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>0.5826</td>
<td>0.7564</td>
<td>0.6155</td>
<td>0.6994</td>
<td>0.6514</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>0.6295</td>
<td>0.8317</td>
<td>0.6832</td>
<td>0.7646</td>
<td>0.7240</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>0.7417</td>
<td>0.8796</td>
<td>0.7683</td>
<td>0.8444</td>
<td>0.8164</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>0.4936</td>
<td>0.8246</td>
<td>0.4955</td>
<td>0.8124</td>
<td>0.5865</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>0.5803</td>
<td>0.7201</td>
<td>0.5947</td>
<td>0.6770</td>
<td>0.6324</td>
<td>10</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 2 contains information about lower and upper bounds of efficiency scores given by modified model (4) - \( \theta_g \) and \( \phi_g \) values are synthesized by (5). Next three columns present similar information from 50 simulation trials as described in previous section of the paper – minimum, maximum and average efficiency characteristics (5). Last two columns of Table 2 compare rankings of DMUs by the middle of interval given by lower and upper bounds from optimization runs on the one side and by average characteristics (5) from simulation approach on the other side. It is clear that both rankings are very close each other. Of course simulation procedure offers much more information to decision makers than single optimization approach. Information about distribution of efficiency scores is not presented here due to a limited space for the paper.

Comparison of results given by two single models for evaluation of efficiency (Table 1) and the two-stage model (Table 2) shows more significant differences in final ranking of DMUs – e.g. DMU_9 is on 5th place in two single models approach and it is one of the worse DMUs when two-stage model is applied.

4 Conclusions

Evaluation of efficiency of network production systems is a very complex task. The paper focused on a simplest system which is two-stage serial model. Under the assumption of deterministic data there are formulated several DEA models for efficiency evaluation. In case of stochastic data one can use optimization approach that offer information about the worse and the best efficiencies (under worse and best conditions for the evaluated unit) only. The same information can be given by simulation approach but except this many other results can be of interest for decision makers. Both approaches are illustrated on a simple numerical example of a real-world nature. Further research will be focused on more complex network systems with serial or parallel structures.
Acknowledgements

The research is supported by the project of the Czech Science Foundation P403/12/1387.

References