

## Time-Simultaneous Prediction Bands: A New Look at the Uncertainty Involved in Forecasting Mortality

Wai-Sum Chan, PhD, CStat, FSA, CERA,  
*Professor of Finance*  
*The Chinese University of Hong Kong*

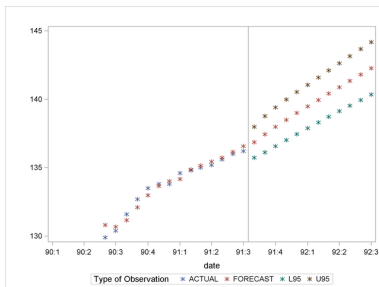
# 1. Introduction

---

- In recent years, there has been a new wave of work that is focused on the forecasting of uncertainty in mortality projections.
- Conventionally, isolated (point-wise) prediction intervals (IPI) are used to quantify the uncertainty in future mortality rates.
- A pointwise interval only reflects uncertainty in a variable at a single time point.
- In situations when the path or trajectory of future mortality rates is important, a band of pointwise intervals might lead to invalid inference.
- The primary objective of this paper is to demonstrate how simultaneous prediction bands (SPB) can be created for prevalent stochastic models.

# What is IPI?

- Most time-series computer packages calculate and plot isolated (point-wise) prediction intervals (IPIs) for multiple forecasts.
- The following example shows an output from a standard forecasting package.
- 95% prediction intervals are included.
- What is the meaning of the confidence level 95% here?



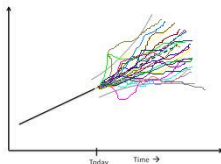
# What is SPB?

---

- SPB is a time-simultaneous prediction band with coverage probability  $0 < 1 - \alpha \leq 1$  for a random trajectory  $\mathbf{y}$  if

$$\Pr(\mathbf{y} \in \mathbf{SPB}) = \Pr\left(\bigcap_{s=1}^S (l_s \leq y_{T+s} \leq h_s)\right) = 1 - \alpha.$$

- SPB is useful when the whole path (or trajectory) of future mortality rates is important.
- The following graph shows a hypothetical 95% SPB.



## Why IPI $\neq$ SPB?

---

- Unless all trajectories develop very orderly (say, perfectly correlated), the probability that a trajectory lies completely inside all ISI would be less than  $1 - \alpha$ .
- It can be seen from an extreme example that if the forecasts are uncorrelated to each other, the probability of  $S$  consecutive predictions are within the ISI is  $(1 - \alpha)^S$ .
- For  $S = 20$  and  $\alpha = 5\%$ , this probability becomes 35.85%!
- In general, predictions from a stochastic mortality model are neither perfectly correlated nor totally uncorrelated; therefore, we cannot use IPI to replace SPB.

## 2. Data

---

- Historic mortality data for US and Canadian (unisex) populations from age 60 to 99 and from year 1951 to 2004.
- The required data, death counts and exposures-to-risk, are obtained from the Human Mortality Database (2010).
- Note that the methods we propose do not require a specific choice of a sample age range and a sample period. They are chosen just purely for illustration purposes.

### 3. Mortality Models

---

- We provide a brief review of two mortality models which we use to illustrate the concept of time-simultaneous prediction bands (SPB).
- They are the Cairns-Blake-Dowd (CBD) model and the generalized CBD model with a cohort effect (GCBD).
- Note that the methods we propose do not require a specific choice of stochastic mortality models. They are chosen just purely for illustration purposes.

## The Cairns-Blake-Dowd (CBD) Model

---

- Cairns et al. (2006) propose a two-factor stochastic mortality model

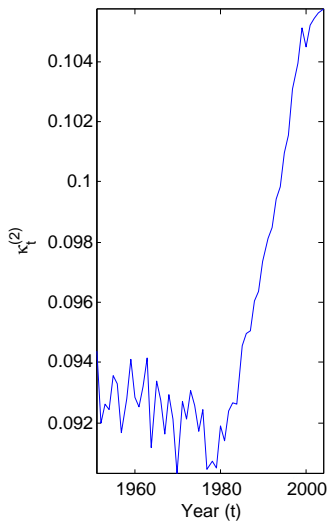
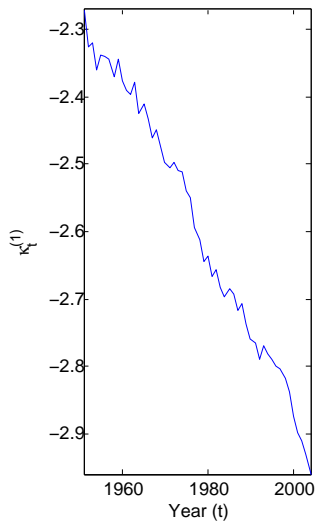
$$\ln \left( \frac{q_{x,t}}{1 - q_{x,t}} \right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}), \quad (1)$$

where  $q_{x,t}$  is the realized single-year death probability at age  $x$  and time  $t$ ,  $\bar{x}$  is the average age over the age range we consider, and  $\kappa_t^{(1)}$  and  $\kappa_t^{(2)}$  are period indexes.

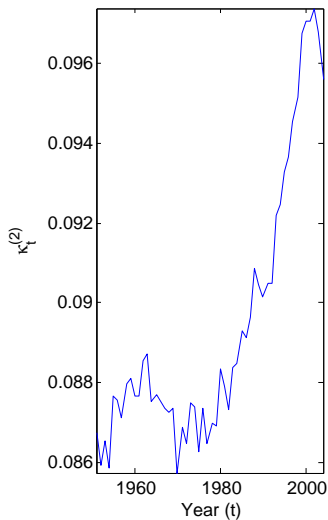
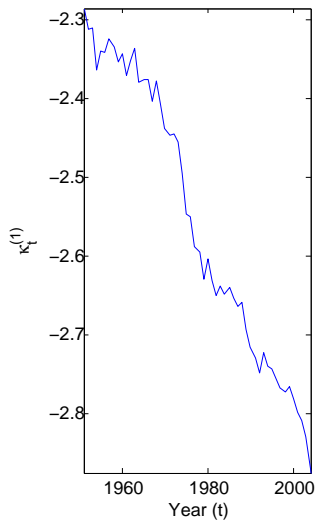
- The maximum likelihood estimates (MLE) of  $\kappa_t^{(1)}$  and  $\kappa_t^{(2)}$ ,  $t = 1950, \dots, 2004$ , for Canada and US, are shown in the following graphs.



## MLE of the CBD Model — Canada



## MLE of the CBD Model — US



## The Cairns-Blake-Dowd (CBD) Model

---

- After fitting equation (1) to historic death probabilities, the period indexes  $\kappa_t^{(1)}$  and  $\kappa_t^{(2)}$  are modeled by a bivariate random walk with drift, that is,

$$\kappa_{t+1} = \kappa_t + \mu + CZ(t+1) \quad (2)$$

where  $\kappa_t = (\kappa_t^{(1)}, \kappa_t^{(2)})'$ ,  $\mu = (\mu_1, \mu_2)'$  is a constant  $2 \times 1$  vector,  $C$  is a constant  $2 \times 2$  upper triangular matrix, and  $Z(t)$  is a 2-dimensional standard normal random vector.

# The Cairns-Blake-Dowd (CBD) Model

---

- Trajectory of mortality predicted rates for a particular birth cohort (age  $x$  in year  $T$ ) can be obtained by

$$\ln \left( \frac{\hat{q}_{x+s, T+s}}{1 - \hat{q}_{x+s, T+s}} \right) = \kappa_T^{(1)}(s) + \kappa_T^{(2)}(s)(x + s - \bar{x}),$$

where  $\kappa_T^{(1)}(s)$  and  $\kappa_T^{(2)}(s)$  are the minimum square error (MMSE) forecasts of  $\kappa_{T+s}^{(1)}$  and  $\kappa_{T+s}^{(2)}$ , respectively.

## The GCBD Model

---

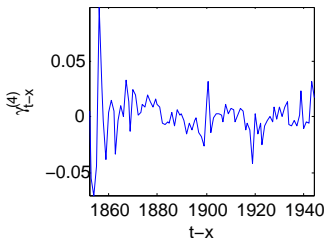
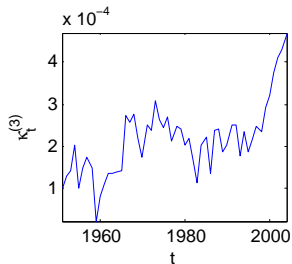
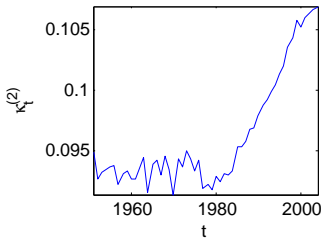
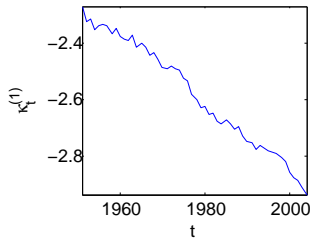
- To model cohort effects, we may consider the following generalization of the CBD model:

$$\ln \left( \frac{q_{x,t}}{1 - q_{x,t}} \right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \kappa_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}^{(4)}, \quad (3)$$

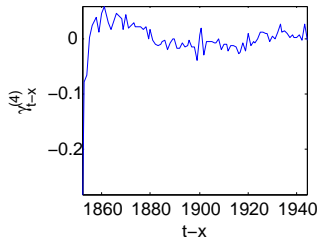
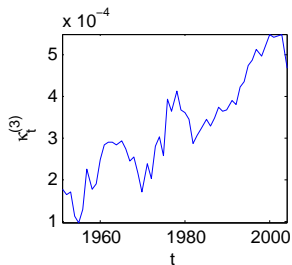
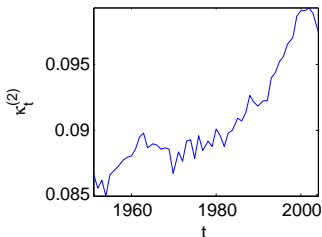
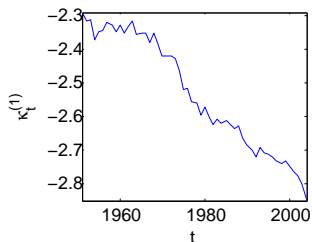
where  $\kappa_t^{(1)}$ ,  $\kappa_t^{(2)}$ , and  $\kappa_t^{(3)}$  are period risk factors,  $\gamma_{t-x}^{(4)}$  is a cohort risk factor, and  $\hat{\sigma}_x^2$  is the mean of  $(x - \bar{x})^2$  over the age range we consider.

- The maximum likelihood estimates (MLE) of  $\kappa_t^{(1)}$ ,  $\kappa_t^{(2)}$ ,  $\kappa_t^{(3)}$  and  $\gamma_{t-x}^{(4)}$ ,  $t = 1950, \dots, 2004$ , for Canada and US, are shown in the following graphs.

# MLE of the GCBD Model — Canada



# MLE of the GCBD Model — US



## The GCBD Model

---

- Having fitted equation (3) to historic data, the period indexes are modeled by a trivariate random walk with drift:

$$\kappa_{t+1} = \kappa_t + \mu + CZ(t+1), \quad (4)$$

where  $\kappa_t = (\kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)})'$ ,  $\mu = (\mu_1, \mu_2, \mu_3)'$  is a constant  $3 \times 1$  vector,  $C$  is a constant  $3 \times 3$  upper triangular matrix, and  $Z(t)$  is a 3-dimensional standard normal random vector.



## The GCBD Model

---

- Trajectory of mortality predicted rates for a particular birth cohort (age  $x$  in year  $T$ ) can be obtained by

$$\ln \left( \frac{\hat{q}_{x+s, T+s}}{1 - \hat{q}_{x+s, T+s}} \right) = \kappa_T^{(1)}(s) + \kappa_T^{(2)}(s)(x + s - \bar{x}) + \kappa_T^{(3)}(s)((x + s - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{T-x}^{(4)}, \quad (5)$$

where  $\kappa_T^{(i)}(s) = \kappa_T^{(i)} + s\mu_i$ ,  $i = 1, 2, 3$ , is the MMSE forecast of  $\kappa_{T+s}^{(i)}$ .

## 4. Simulation-Based SPB

---

- Consider the cohort of individuals who are aged  $x$  at the forecast origin  $T$ . Assuming that the forecast horizon is  $S$ , the trajectory of interest would be  $\mathbf{m} = (m_{x+1, T+1}, \dots, m_{x+S, T+S})$ . From the stochastic components of the mortality models, we simulate a **learning sample** of size  $N$  (we use  $N = 5000$  in this paper):

$$\mathbf{M} = \{\mathbf{m}^{(n)}\}_{n=1}^N = \{(m_{x+1, T+1}^{(n)}, \dots, m_{x+S, T+S}^{(n)})\}_{n=1}^N$$

- Note that CBD models is based on  $q_{x,t}$  rather than  $m_{x,t}$ . To obtain central death rates, we apply the constant force of mortality (CFM) assumption:

$$m_{x,t} = -\ln(1 - q_{x,t}),$$

## Method 1: Adjusted Intervals

---

- Our goal is to construct a time-simultaneous prediction band such that  $\lceil(1 - \alpha)N\rceil$  of the trajectories in  $\mathbf{M}$  are completely inside the band at every time point, while  $\lfloor\alpha N\rfloor$  are outside the band at one or more points of time.
- Let  $PI_s = [l_s, h_s]$  be a pointwise prediction interval for  $m_{x+s, T+s}$  with coverage probability  $1 - \alpha$ . Then the limits  $l_s$  and  $h_s$  are set to the  $\lfloor\alpha/2\rfloor$ th lowest and highest values in the sample, respectively.
- Kolsrud (2007) proposes a three-step iterative method, which he calls 'adjusted intervals,' to construct a time-simultaneous prediction band from a learning sample.

## Method 1: Adjusted Intervals

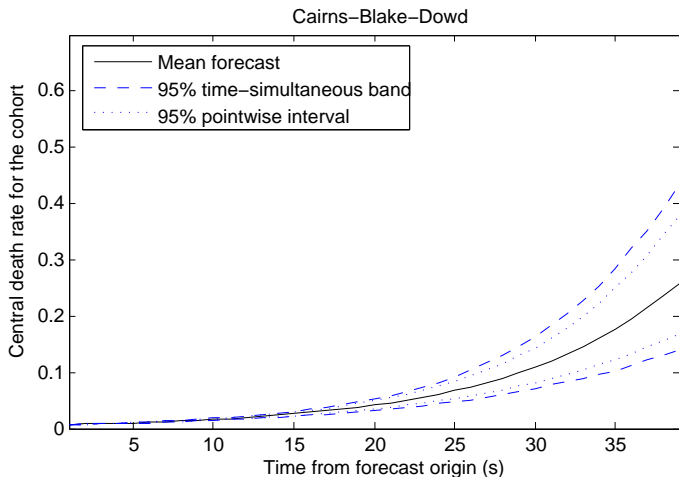
---

- The iterative algorithm:
  - **Step 1:** For each  $s = 1, \dots, S$ , widen the interval uniformly to include the nearest sample point above and the nearest sample point below.

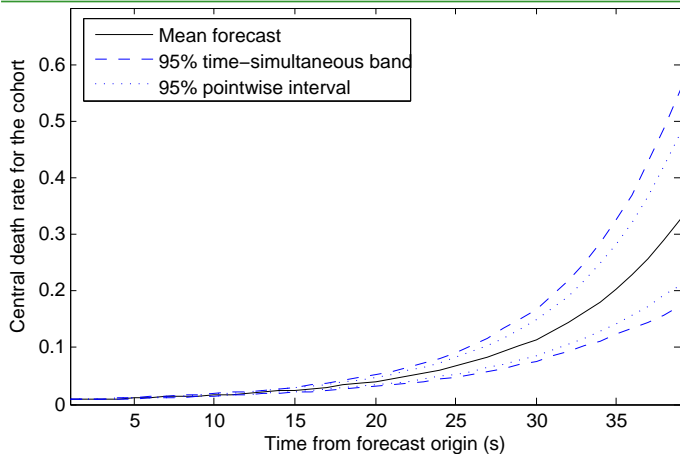
$$PI_s^* = [l_s - \delta_s, h_s + \delta_s]$$

- **Step 2:** Check the simultaneous coverage of all intervals in the learning sample  $\mathbf{M}$ .
  - **Step 3:** If the simultaneous coverage is less than the prescribed level  $1 - \alpha$ , go to Step (1). Otherwise, terminate the algorithm.
- The resulting band of intervals would contain no less than  $1 - \alpha$  of the trajectories in the learning sample.

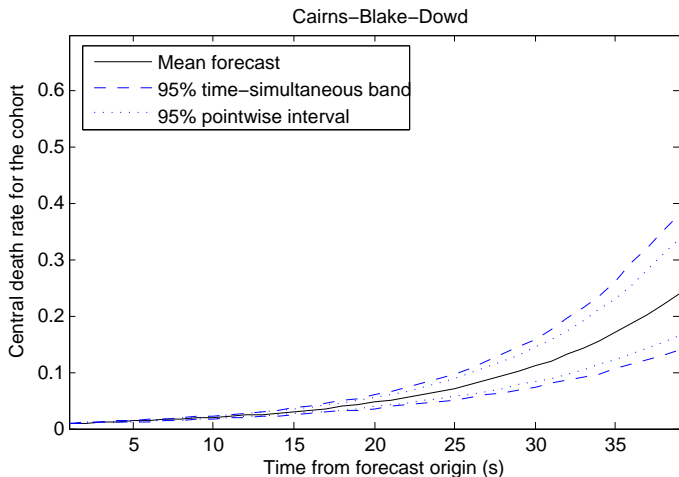
# Illustration of Adjusted Intervals — Canada



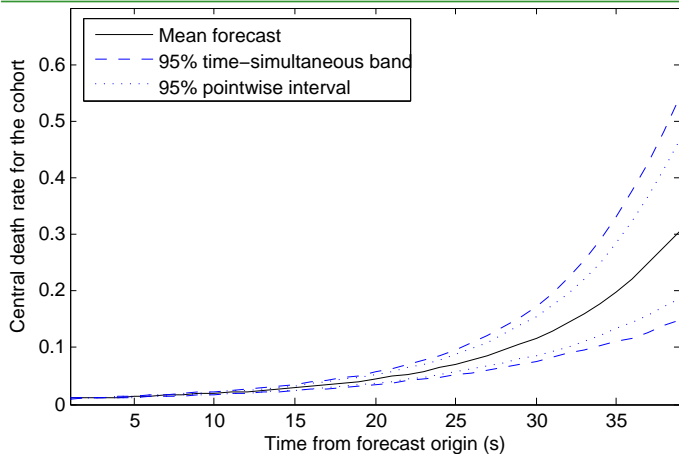
## Generalized Cairns–Blake–Dowd



# Illustration of Adjusted Intervals — US



## Generalized Cairns–Blake–Dowd





## Method 2: Chebyshev Bands

---

- The **envelope** of a (sub)sample is defined as the tightest band that contains all trajectories in the (sub)sample.
- As an example, the envelope of the learning sample  $\mathbf{M}$  can be expressed as  $([\min_n m_{x+s, T+s}^{(n)}, \max_n m_{x+s, T+s}^{(n)}])_{s=1}^S$ .
- The idea behind Chebyshev bands is that we construct a time-simultaneous prediction band as the envelope of a subsample  $\mathbf{M}^*$  that contains  $\lceil (1 - \alpha)N \rceil$  trajectories with the shortest distance to the mean trajectory

$$\bar{\mathbf{m}} = (\bar{m}_{x+1, T+1}, \dots, \bar{m}_{x+S, T+S}).$$

## Method 2: Chebyshev Bands

---

- It is suggested that we measure the distant by the the weighted Chebyshev distance, which can be expressed as:

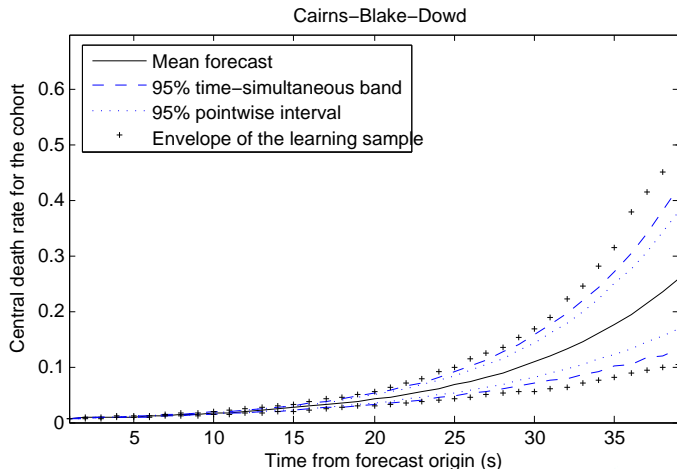
$$\mathcal{D} = \max_{s=1,\dots,S} \left( \frac{|m_{x+s,T+s} - \bar{m}_{x+s,T+s}|}{\sigma_s} \right),$$

where

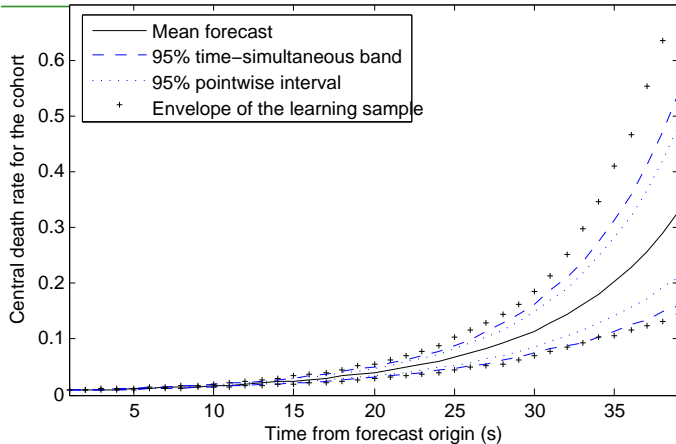
$$\sigma_s = \sqrt{\frac{1}{N} \sum_{n=1}^N (m_{x+s,T+s} - \bar{m}_{x+s,T+s})^2}$$

is the pointwise standard deviation  $s$  steps beyond the forecast origin.

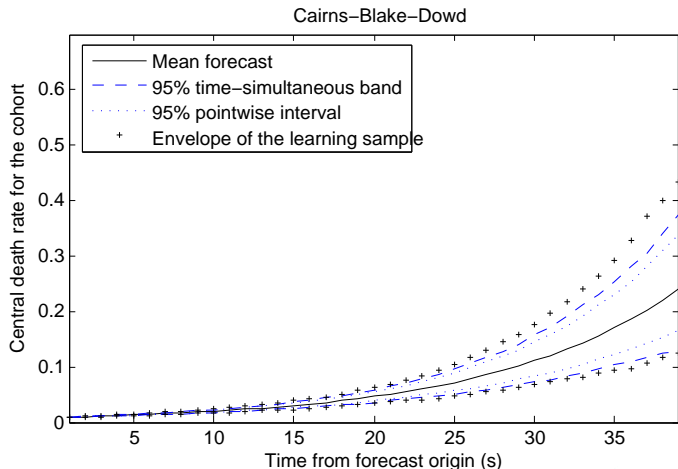
# Illustration of Chebyshev Bands — Canada



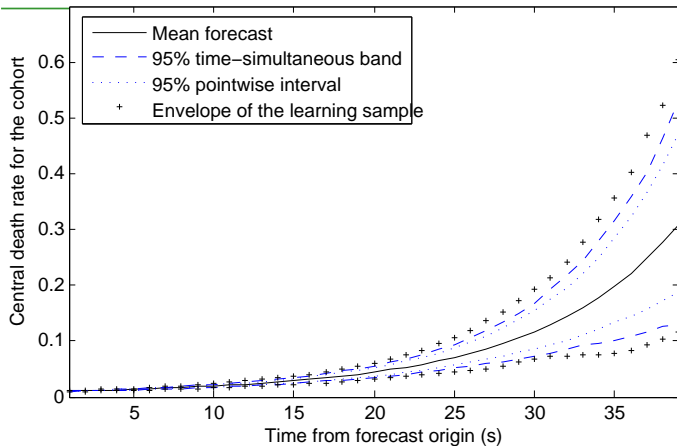
Generalized Cairns–Blake–Dowd



# Illustration of Chebyshev Bands — US



Generalized Cairns–Blake–Dowd



## 5. Further Research

---

- Model Risk: Denuit (2009)
- Parameter Method: Li and Chan (2011)
- Non-linear Model: Azaïs et al. (2010)

Thank You!