COHERENT INCURRED PAID (CIP) MODELS FOR CLAIMS RESERVING

Gilles DUPIN, (1) Emmanuel KOENIG, (1) Pierre LE MOINE, (1)
Alain MONFORT (2), E. RATIARISON (1)

August 2015

[1]Groupe MONCEAU, Paris
[2]CREST Paris
Corresponding author : Alain Monfort, monfort@ensae.fr

COHERENT INCURRED PAID (CIP)MODELS FOR CLAIMS RESERVING

Abstract

In this paper we first propose a statistical model, called Coherent Incurred Paid (CIP) model, to predict future claims, using simultaneously the information contained in incurred and paid claims. This model does not assume log-normality of the levels (or normality of the growth rates) and is semi-parametric since it only specifies the first and the second moments. Correlations between growth rates of incurred and paid claims are allowed and the tail development period is estimated. We also propose methods for computing the Claim Development Results (CDR) and their Values at Risk (VaR) in this semi-parametric framework. Moreover we show how to take into account the updating of the estimation in the computation of the CDR's. An application highlights the practical importance of relaxing the normality assumption and of updating the estimation of the parameters.

Keywords : Incurred and paid claims, simultaneous estimation, correlation, semi-parametric approach, tail development, prediction, updating, CDR, VaR.

JEL : C14, C15, C51, C52, C53.

1 INTRODUCTION

One of the more important problems that non-life insurance companies have to solve is the evaluation of the reserve risk. Such an evaluation necessitates a two step modelling. The first step requires a method of prediction of the ultimate claims. The second one has to define a measure of the reserve risk based on these predictions. Let us consider more precisely these two steps.

There exists a large literature dealing with the first step. Most methods are based either on cumulated payments or on incurred losses. However there are important works proposing models using both sources of information. Halliwell (1997, 2009) and Venter (2008) used a regression approach. Quang and Mack (2004) introduced the Munich Chain Ladder involving a modification the chain Ladder development factors based on incurred-paid ratios. Postuma, Cator, Veerkamp and van Zwet (2008) suggested a multivariate model conditioned on equality of the total paid and incurred losses. Merz and Wuthrich (2010a) proposed a probabilistic model, the Paid Incurred Chain (PIC) model, combining in a rigorous way the two kinds of information. This work has been followed by several extensions incorporating new features : tail development factors [Merz and Wuthrich (2010b)], claim development results [Happ, Merz and Wuthrich (2011)]] or dependence [Happ and Wuthrich (2011)].

The second step of the modelling is the definition and the computation of the reserve risk. Following recommendations of regulatory authorities, the more popular measure is based on the Claim Development Result (CDR) defined as the difference between the prediction of the ultimate claims to day and in one year time. More precisely the measure of the reserve risk is the $99,5\%$ quantile of the opposite of the CDR, viewed as evaluating under-provisioning : this is the so-called Value at Risk (VaR) notion. [see Wuthrich and Merz (2008)].

In the present paper we consider both steps of the modelling strategy. We propose a statistical method, the Coherent Incurred Paid (CIP) method, using simultaneously information based on paid and incurred claims. This method may be parametric or semi-parametric. In the parametric case we make a Gaussian assumption and we use a conditional approach in the spirit of Posthama et al. (2008) ; in the semi-parametric case, we only propose parametric specifications of the expectations and of the variance-covariance matrix of the bidimensional vector composed of the rates of increase of the cumulated payments and incurred losses, and the whole distribution is estimated by non-parametric kernel methods.

This approach has several advantages. First, since the number of parameters do not increase with the number of observations, we can use the standard asymptotic theory of the Maximum Likelihood (ML) method (in the parametric case) or the Pseudo-Maximum Likelihood (PML) method (in the semi-parametric case) and, in particular, we can test the significativity of the estimators. Second, this approach is flexible enough to incorporate a correlation between the two sources of informations and to estimate the date at which predictions of the ultimate claims based on both kinds of information become equal. Third, the computation of risk measures, namely the VaR's based on the CDR's, can be made with or without the normality assumption and with or without incorporating the uncertainty on the estimation of the parameters. An application on incurred claims and cumulated payments corresponding to a line of business Motor Body Liability-Insurance highlights several points. First, our CIP method is easily implementable. Second the projected values of the incurred claims and cumulated payments corresponding to the largest observed development years are very different. Third these values are also very different from the ones provided by the Chain Ladder method, and the CIP method provides a unique ultimate value which is located between the ultimate values of the Chain Ladder method. Fourth in the computation of the VaR's of the CDR's it is crucial to take into account the non-gaussianity of the rates of increase and the updating of the estimations.

The paper is organized as follows. Section 2 describes the Gaussian CIP model, the estimation of its parameters and of the tail development year, the computation of prediction as well as CDR's and their VaR's. Section 3 generalizes these results to a semi-parametric framework in which Gaussianity is no longer assumed. Section 4 proposes an application. Section 5 provides concluding remarks. Proofs and data are gathered in appendices.

2 A GAUSSIAN CIP MODEL

2.1 Notations

We denote respectively by $P_{i,j}$ and $I_{i,j}$, the cumulated payments and incurred losses for accident year i and development year j. The calendar year is $i+j$. We will also use the following notations :

$$\begin{aligned}
&X_{1,i,j} = Log\, P_{i,j} \\
&X_{2,i,j} = Log\, I_{i,j} \\
&X_{i,j} = \begin{pmatrix} X_{1,i,j} \\ X_{2,i,j} \end{pmatrix} \\
&Y_{1,i,j} = X_{1,i,j} - X_{1,i,j-1} = Log\, \tfrac{P_{i,j}}{P_{i,j-1}} \\
&Y_{2,i,j} = X_{2,i,j} - X_{2,i,j-1} = Log\, \tfrac{I_{i,j}}{I_{i,j-1}} \\
&Y_{i,j} = \begin{pmatrix} Y_{1,i,j} \\ Y_{2,i,j} \end{pmatrix}
\end{aligned}$$

We assume that $X_{i,j}$ is observed for :

$$\begin{aligned}
&i = 1, ..., n \\
&j = 0, ..., n-1 \\
&1 \leq i+j \leq n
\end{aligned}$$

and, consequently, $Y_{i,j}$ is observed for :

$$\begin{aligned}
&i = 1, ..., n-1 \\
&j = 1, ..., n-1 \\
&2 \leq i+j \leq n
\end{aligned}$$

We also introduce the notations :

$$\begin{array}{lll}
Y_i = (Y'_{i,1}, ..., Y'_{i,n-i})' & \text{of size} & 2(n-i) \\
Y_{1,i} = (Y_{1,i,1}, ..., Y_{1,i,n-i})' & \text{of size} & n-i \\
Y_{2,i} = (Y_{2,i,1}, ..., Y_{2,i,n-i})' & \text{of size} & n-i
\end{array}$$

A key assumption, throughout the paper, is that there is an ultimate development year $N \geq n-1$, in general not observed, such that $X_{1,i,N} = X_{2,i,N}$ for all i, and the models proposed will have to satisfy this constraint. We also introduce the notations :

$$\begin{array}{lll} \widetilde{Y}_i = (Y'_{i,1}, ..., Y'_{i,N})' & \text{of size} & 2N \\ \widetilde{Y}_{1,i} = (Y_{1,i,1}, ..., Y_{1,i,N})' & \text{of size} & N \\ \widetilde{Y}_{2,i} = (Y_{2,i,1}, ..., Y_{2,i,N})' & \text{of size} & N \end{array}$$

2.2 A conditional Gaussian model

A first Coherent Incurred Paid (CIP) model is obtained by starting from a Gaussian model and imposing the conditioning constraints :

$$X_{1,i,N} = X_{2,i,N}, i = 1, ..., n-1.$$

More precisely we assume that $X_{i,0}$ is fixed and that :

$$Y_{i,j} = m(i,j,\theta) + \xi_{i,j} \tag{1}$$

where the $m(i,j,\theta)$ are bidimensional deterministic functions and the $\xi_{i,j}$ are bidimensional vectors following independently the Gaussian distribution $N[0, \Omega(i,j,\theta)]$, where θ is an unknown vector of parameters. Note that $\Omega(i,j,\theta)$ is not assumed to be diagonal and, therefore, a correlation between $Y_{1,i,j}$ and $Y_{2,i,j}$ is allowed.

Then we impose the new information :

$$X_{1,i,N} = X_{2,i,N}, i = 1, ..., n-1 \tag{2}$$

In other words we assume that the $\widetilde{Y}_i, i = 1, ..., n-1$ are independently distributed and that the distribution of $\widetilde{Y}_i$ is the conditional distribution obtained from the initial Gaussian model (1) by imposing $X_{1,i,N} = X_{2,i,N}$ or, equivalently, $d'X_{i,N} = 0$ with $d' = (1, -1)$

Proposition 1

The conditional distribution of $\widetilde{Y}_i$ given $d'X_{i,N} = 0$ is the Gaussian distribution :

$$N\left(\widetilde{m}_i - \frac{\widetilde{c}_i a_i}{b_i}, \widetilde{\Omega}_i - \frac{\widetilde{c}_i \widetilde{c}_i'}{b_i}\right)$$

where :

$$\widetilde{m}_i = [m'(i,1,\theta), ..., m'(i,N,\theta)]'$$

$$\widetilde{c}_i = [d'\Omega(i,1,\theta), ..., d'\Omega(i,N,\theta)]'$$

$$\widetilde{\Omega}_i = \begin{bmatrix} \Omega(i,1,\theta) & & 0 \\ & \dots & \\ 0 & & \Omega(i,N,\theta) \end{bmatrix}$$

$$a_i = d'X_{i,0} + d'\sum_{j=1}^{N} m(i,j,\theta)$$

$$b_i = d'\sum_{j=1}^{N} \Omega(i,j,\theta)d$$

Proof : see appendix 1

Note that using the notation $F_N = (I_2, ..., I_2)'$ where the identity matrix of size 2 is repeated N times, we have

$$\widetilde{c}_i = \widetilde{\Omega}_i F_N d$$

$$a_i = d'(X_{i,0} + F_N'\widetilde{m})$$

$$b_i = d'F_N'\widetilde{\Omega}_i F_N d = d'F_N'\widetilde{c}_i$$

In particular we deduce the conditional distribution of the observed vector Y_i.

Corollary 1
The conditional distribution of Y_i given $d'X_{i,N} = 0$ is the Gaussian distribution :

$$N\left(m_i - \frac{c_i a_i}{b_i}, \Omega_i - \frac{c_i c_i'}{b_i}\right)$$

where :

$$m_i = [m'(i,1,\theta), ..., m'(i,n-i,\theta)]'$$
$$c_i = [d'\Omega(i,1,\theta), ..., d'\Omega(i,n-i,\theta)]'$$
$$\Omega_i = \begin{bmatrix} \Omega(i,1,\theta) & & 0 \\ & \cdots & \\ 0 & & \Omega(i,n-i,\theta) \end{bmatrix}$$

Proof : we just have to take the marginal distribution of the first $n-i$ components of the joint distribution given in proposition 1. $\square$

Note that the Gaussian distribution of proposition 1 is degenerated since the components of $\widetilde{Y}_i$ have to satisfy the linear constraint :

$$d'X_{i,0} + d'\sum_{j=1}^{N} Y_{i,j} = 0$$

or :

$$d'(X_{i,o} + F_N'\widetilde{Y}_i) = 0.$$

The matrix :

$$\widetilde{\Omega}_i - \frac{\widetilde{c}_i \widetilde{c}_i'}{b_i}$$

is of rank $2N-1$. However as soon as n is strictly smaller than $N+1$ the variance-covariance matrix of Y_i namely :

$$\Omega_i - \frac{c_i c_i'}{b_i}$$

is of full rank $2(n-i)$ for all i including $i = 1$.

2.3 Estimation of a Gaussian CIP model

As soon as the functions $m(i,j,\theta)$ and $\Omega(i,j)$ have been specified (see section 4 for a discussion of these specifications) the parameter θ can be estimated by the Maximum Likelihood (ML) method. Indeed from corollary 1 we deduce that the log-likelihood function of the model is :

Proposition 2

$$L_n(\theta) = -\frac{1}{2}\sum_{i=1}^{n-1}\left[Logdet\Sigma_i(\theta) + (y_i - \mu_i(\theta))'\Sigma_i^{-1}(\theta)(y_i - \mu_i(\theta))\right]$$

with :

$$\mu_i(\theta) = m_i(\theta) - \frac{c_i(\theta)a_i(\theta)}{b_i(\theta)}$$

$$\Sigma_i(\theta) = \Omega_i(\theta) - \frac{c_i(\theta)c_i'(\theta)}{b_i(\theta)}$$

Proof : it is a direct consequence of the expression of the probability density function of a multivariate Gaussian distribution. □

Moreover the computation of $\Sigma_i^{-1}(\theta)$ is simple thanks to the following proposition (omitting θ for notational simplicity).

Proposition 3

$$\Sigma_i^{-1} = \Omega_i^{-1} + \frac{\Omega_i^{-1}c_ic_i'\Omega_i^{-1}}{b_i - c_i'\Omega_i^{-1}c_i}$$

with :

$$\Omega_i^{-1} = \begin{pmatrix} \Omega_{i,1}^{-1} & & 0 \\ & \dots & \\ 0 & & \Omega_{i,n-i}^{-1} \end{pmatrix}$$

Proof : see appendix 2

In particular the previous proposition implies that the term $(y_i-\mu_i)'\Sigma_i^{-1}(y_i-\mu_i)$ in the log-likelihood is simply :

$$\sum_{j=1}^{n-i}\left[(y_{i,j}-\mu_{i,j})'\Omega_{i,j}^{-1}(y_{i,j}-\mu_{i,j})+\frac{[(y_{i,j}-\mu_{i,j})'\Omega_{i,j}^{-1}c_{i,j}]^2}{b_i-\sum_{j=1}^{n-i}c'_{i,j}\Omega_{i,j}^{-1}c_{i,j}}\right]$$

Starting values for the parameters appearing in the $m(i,j,\theta)$ can be obtained by the ordinary least squares (OLS) method and from the residuals of this method for the parameters appearing in the $\Omega(i,j,\theta)$ [see section 4]

The ML estimator of θ will be denoted by $\widehat{\theta}_n$. Note that it is based on $(n-1)+(n-2)+...+1=\frac{n(n-1)}{2}$ observations $Y_{i,j}$ of size 2.

The whole testing and confidence region methods based on ML estimators apply. In particular the variance-covariance matrix of $\widehat{\theta}_n$ can be approximated by :

$$-\left[\frac{\partial^2 L_n(\widehat{\theta}_n)}{\partial\theta\partial\theta'}\right]^{-1}$$

2.4 Tail development

As mentioned in section 2.1, the ultimate development year N is, in general, larger than the latest development year j where observations of Y_i are available, namely $j=n-1$. It is the so called "tail development" problem. In the context of chain ladder approaches, this problem has been reduced to the computation of an ultimate development factor called "tail development factor". A review of there methods is available in Boor (2006). The tail development problem has also been considered by Merz and Wuthrich (2010) within their PIC method; in particular their bayesian approach allows for a tail development factor covering several development periods beyond the last column of the claim development triangle.

In our approach we consider N as an unknow parameter. It is clear that the previous log-likelihood depends on N through the $a_i's$ and the $b_i's$. Therefore we can use recent results on the estimation of discrete parameter models (see Choirat and Seri (2012)) showing that maximizing the log-likelihood function with respect to all the parameters, including N, provides a consistent estimator of N. This gives us not only coherent estimates of the ultimate values $P_{i,N} = I_{i,N}$, for any i, but also an estimate of the length of the tail development period.

2.5 Prediction

Once the parameters are estimated, we have to predict $Y_i^* = (Y'_{i,n-i+1}, ..., Y'_{i,N})'$, or, equivalently, $X_i^* = (X'_{i,n-i+1}, ..., X'_{i,N})'$ for each i, given the observations $X_{i,0}, Y_{i,1}, ..., Y_{i,n-i}$ or, equivalently $X_i = (X'_{i,0}, ..., X'_{i,n-i})'$.

The conditional distribution of X_i^* given X_i (without conditioning by $d'X_{i,N} = 0$) is the same as the conditional distribution of X_i^* given $X_{i,n-i}$ since for any $k \in \{1, ..., N-n+i\}$:

$$X_{n-i+k} = X_{i,n-i} + \sum_{j=1}^{k} Y_{i,n-i+j}$$

and the $Y_{i,n-i+k}$ are independent of X_i. Therefore the conditional distribution of X_i^* given X_i and $d'X_{i,N} = 0$ is the same as the conditional distribution of X_i^* given $X_{i,n-i}$ and $d'X_{i,N} = 0$ (since $d'X_{i,N}$ is function of X_i^*).

This implies that we have to solve the same problem as in section 2.2, just replacing $X_{i,0}$ by $X_{i,n-i}$ and $\widetilde{Y}_i$ by Y_i^*, and we get the following proposition.

Proposition 4

The conditional distribution of Y_i^* given $X_{i,n-i}$ and $d'X_{i,N} = 0$ is the Gaussian distribution :

$$N\left(m_i^* - \frac{c_i^* a_i^*}{b_i^*}, \Omega_i^* - \frac{c_i^* c_i'^*}{b_i^*}\right)$$

with :

$$m_i^* = [m'(i, n-i+1, \theta), ..., m'(i, N, \theta)]'$$

$$c_i^* = [d'\Omega(i, n-i+1, \theta), ..., d'(\Omega_i, N, \theta)]'$$

$$\Omega_i^* = \begin{bmatrix} \Omega(i, n-i+1, \theta) & & 0 \\ & \cdots & \\ 0 & & \Omega(i, N, \theta) \end{bmatrix}$$

$$a_i^* = d'\left(X_{i,n-i} + \sum_{j=n-i+1}^{N} m(i, j, \theta)\right)$$

$$b_i^* = d' \sum_{j=n-i+1}^{N} \Omega(i, j, \theta) d$$

We also will also use the notations :

$$\mu_i^* = m_i^* - \frac{c_i^* a_i^*}{b_i^*}, \Sigma_i^* = \Omega_i^* - \frac{c_i^* c_i'^*}{b_i^*}$$

$$\text{and } \mu_i^* = [\mu^{*'}(i, n-i+1, \theta), \ldots, \mu^{*'}(i, N, \theta)]'$$

The best prediction of Y_i^* is μ_i^* and the best prediction of X_{n-i+k}, $k = \{1, \ldots, N-n+i\}$ is $X_{i,n-i} + \sum_{j=1}^{k} \mu^*(i, n-i+j, \theta)$

2.6 Claim Development Results (CDR)

Denoting by E_n the estimation of the conditional expectation operator with respect to the true conditional distribution given the information at the calendar date n :

$$\mathcal{J}_n = \{X_{i,0}, i = 1, ..., n-1, Y_{i,j}, i = 1, ..., n-1, j = 1, ..., n-i\}$$

evaluated at $\widehat{\theta}_n$, the Claim Development Result for the accounting calendar period $(n, n+1)$ and accident year i is :

$$CDR_i(n+1) = E_n(X_{1,i,N}) - E_{n+1}(X_{1,i,N}) \tag{3}$$

or, equivalently :

$$CDR_i(n+1) = E_n(X_{2,i,N}) - E_{n+1}(X_{2,i,N}) \tag{4}$$

since in our model we automatically have $X_{1,i,N} = X_{2,i,N}$.
Choosing $X_{1,i,N}$, we can write :

$$X_{1,i,N} = X_{1,i,n-i} exp\left(\sum_{j=n+1-i}^{N} Y_{1,i,j}\right) = X_{1,i,n-i} exp(f_i' Y_i^*) \tag{5}$$

Where f_i' is the row vector of size $2\,(N-n+i)$ equal to $(1,0,1,0,...,1,0)$ picking the components $Y_{1,i,j}, j = n-i+1, ..., N$ in Y_i^*.

Therefore the true conditional expectation (evaluated at θ_0) of $X_{1,i,N}$ is :

$$X_{1,i,n-i}\, exp[f_i'\mu_i^*(\theta_0) + 1/2 f_i'\Sigma_i^*(\theta_0) f_i] \tag{6}$$

where θ_0 is the true value of θ.

Replacing θ by the maximum likelihood estimator $\widehat{\theta}_n$ we get :

$$E_n(X_{1,i,N}) = X_{1,i,n-i} exp\left[f_i'\mu_i^*(\widehat{\theta}_n) + 1/2 f_i'\Sigma_i^*(\widehat{\theta}_n) f_i\right] \tag{7}$$

Similarly we have :

$$E_{n+1}(X_{1,i,N}) = X_{1,i,n-i+1} exp\left[f_i'^*\mu_i^{**}(\widehat{\theta}_{n+1}) + 1/2 f_i'^*\Sigma_i^{**}(\widehat{\theta}_{n+1}) f_i^*\right] \tag{8}$$

where f_i^* and μ_i^{**} are obtained from f_i and μ_i^*, respectively, by deleting the first two components and Σ_i^{**} is obtained from Σ_i^* by deleting the first two rows and the first two columns.

$X_{1,i,n-i+1}$ is random at date n and is equal to $X_{1,i,n-i} exp(Y_{1,i,n-i+1})$ with :

$$Y_{1,i,n-i+1} = \mu_{1,i}^*(\theta_0^*) + \sigma_{1,i}^*(\theta_0)\varepsilon_{1,i,n-i+1}$$

where $\mu^*_{1,i}(\theta_0)$ is the first component of $\mu^*_i(\theta_0), \sigma^*_{1,i}(\theta_0)$ the square root of the $(1,1)$ term of $\Sigma^*_i(\theta_0)$ and $\varepsilon_{1,i,n-i+1}$ is following $N(0,1)$.

It is natural to view $CDR_i(n+1)$ from the calendar date n and, therefore, to replace θ_0 by $\widehat{\theta}_n$ in the expression above of $Y_{1,i,n-i+1}$.

Finally we get the evaluation of $CDR_i(n+1)$:

$$\widehat{CDR}_i(n+1) = X_{1,i,n-i}\left\{exp\left[f_i'\mu^*_i(\widehat{\theta}_n) + \frac{1}{2}f_i'\Sigma^*_i(\widehat{\theta}_n)f_i\right]\right. \quad (9)$$

$$\left. -exp[\mu^*_{1,i}(\widehat{\theta}_n) + \sigma^*_{1,i}(\widehat{\theta}_n)\varepsilon_{1,i,n-i+1} + f_i'^{*}\mu^{**}_i(\widehat{\theta}_{n+1}) + 1/2 f_i'^{*}\Sigma^{**}_i(\widehat{\theta}_{n+1})f^*_i]\right\}$$

At date $n, \widehat{CDR}_i(n+1)$ is random through $\varepsilon_{1,i,n-i+1}$ and through the $\varepsilon_{1,k,n-k+1}, k = 1, ..., n-1$ (containing $\varepsilon_{1,i,n-i+1}$) and $\varepsilon_{2,k,n-k+1}$ appearing in the new observations $Y_{i,n-i+1}$ at calendar date $n+1$ which are used in the estimation of $\hat{\theta}_{n+1}$. The global CDR is approximated by $\widehat{CDR}(n+1) = \sum_{i=1}^{n}\widehat{CDR}_i(n+1)$.

2.7 Value at Risk (VaR) of the $\widehat{CDR}_i(n+1)$

The Value at Risk $VaR_i(\alpha)$ associated with $\widehat{CDR}_i(n+1)$, or rather with the under-provisioning measure $-\widehat{CDR}_i(n+1)$, is defined by :

$$P[-\widehat{CDR}_i(n+1) < VaR_i(\alpha)] = \alpha$$

where α is close to 1, for instance 0.995.

If we do not take into account the updating of $\widehat{\theta}_n$ and set $\widehat{\theta}_{n+1} = \widehat{\theta}_n$, the only random term in (9) is $\varepsilon_{1,i,n-i+1}$ distributed as $N(0,1)$. In other words, with obvious notations,$\widehat{CDR}_i(n+1)$ is of the form $\beta_i - \gamma_i exp(\delta_i\varepsilon_{1,i,n-i+1})$ with $\beta_i > 0, \gamma_i > 0, \delta_i > 0$.

The $VaR_i(\alpha)$ is easily seen to be :

$$\gamma_i exp[\delta_i\Phi(\alpha)] - \beta_i \quad (10)$$

If we want to take into account the updating of $\widehat{\theta}_n$ into $\widehat{\theta}_{n+1}$, we might use the Newton-Raphson approximation :

$$\widehat{\theta}_{n+1} = \widehat{\theta}_n - \left[\frac{\partial^2 L_n(\widehat{\theta}_n)}{\partial\theta\partial\theta'}\right]^{-1} \frac{\partial L_{n+1}}{\partial\theta}(\widehat{\theta}_n) \tag{11}$$

where $\frac{\partial^2 L_n}{\partial\theta\partial\theta'}(\widehat{\theta}_n)$ is a by-product of the estimation procedure and $\frac{\partial L_{n+1}}{\partial\theta}(\widehat{\theta}_n)$ can be computed numerically as a function of the $\varepsilon_{k,n-k+1}$, $k = 1, ..., n-1$.

Then $VaR_i(\alpha)$ can be evaluated by simulation. More precisely let us consider M simulations of $\widehat{CDR}_i(n+1)$ and let us order them in increasing order, then $-VaR_i(\alpha)$ is taken equal to the value with index $[M\alpha]$ (where [.] is a notation for the integer).

Note that for the computation of the $VaR(\alpha)$ of the global $\widehat{CDR}(n+1) = \sum_{i=1}^n \widehat{CDR}_i(n+1)$ such a simulation method is required even when we do not update $\widehat{\theta}_n$.

3 A GENERAL CLASS OF CIP MODELS

3.1 Semi-parametric models

In the Gaussian model the vectors Y_i, of size $2(n-i)$, follow independently the distribution :

$$N(\mu_i, \Sigma_i)$$

with :

$$\mu_i = m_i - \frac{c_i a_i}{b_i}, \Sigma_i = \Omega_i - \frac{c_i c_i'}{b_i}$$

(see Corollary 1)

Let us denote by T_i the lower triangular matrix, or Cholesky matrix, such that $\Sigma_i = T_i T_i'$ (imposing positive diagonal terms for T_i implies its uniqueness) we can write :

$$Y_i = \mu_i + T_i\varepsilon_i \tag{12}$$

$$\varepsilon_i \sim N(0, I_{2(n-i)}) \quad \text{or}$$

$$\varepsilon_{i,k} \sim IIN(0,1), \;\; k = 1, ..., 2(n-i)$$

A natural extension of this model consists in still assuming $Y_i = \mu_i + T_i\varepsilon_i$ but only imposing that the $\varepsilon_{i,k}$ are identically, independently distributed with zero mean and unit variance :

$$\begin{aligned} Y_i &= \mu_i + T_i\varepsilon_i \\ \varepsilon_{i,k} &\sim II(0,1) \end{aligned} \tag{13}$$

In other words we no longer assume that the $\varepsilon_{i,k}$ are Gaussian and we do not make any assumption about their common distribution. The model becomes semi-parametric and the maximum likelihood method is no longer available.

3.2 Estimation

It can be shown [see Gourieroux-Monfort-Trognon (1984)] that if we estimate θ_0 by the pseudo ML estimator obtained by maximizing $L_n(\theta)$ given in proposition we still obtain a consistent and asymptotically Gaussian estimators for any true distribution of the $\varepsilon'_{i,k}s$.

We can then estimate the ε_i by :

$$\widehat{\varepsilon}_i = \widehat{T}_i^{-1}(Y_i - \widehat{\mu}_i) \tag{14}$$

where $\widehat{T}_i$ and $\widehat{\mu}_i$ are T_i and μ_i evaluated at $\widehat{\theta}_n$.

The unknown distribution of the $\varepsilon'_{i,k}s$ can be estimated by the Gaussian kernel method and we get the following mixture of Gaussian distributions :

$$\frac{1}{n(n-1)h} \sum_{i=1}^{n-1} \sum_{k=1}^{2(n-i)} \varphi\left(\frac{\varepsilon - \widehat{\varepsilon}_{i,k}}{h}\right)$$

where φ is the p.d.f of $N(0,1)$ and h is equal to $[n(n-1)]^{-1/5}$ according to Silverman's rule. The mean and the variance of this mixture of Gaussian distributions are :

$$\overline{\varepsilon} = \frac{1}{n(n-1)} \Sigma_{i,k} \widehat{\varepsilon}_{i,k}$$

$$\overline{\sigma}^2 = h^2 + \frac{1}{n(n-1)}\Sigma_{i,k}\widehat{\varepsilon}^2_{i,k} - \overline{\varepsilon}^2$$

and, in order to get a distribution which is exactly zero mean and unit variance we can use :

$$\frac{\overline{\sigma}}{n(n-1)h}\Sigma_{i,k}\varphi\left(\frac{\overline{\sigma}\varepsilon + \overline{\varepsilon} - \widehat{\varepsilon}_{i,k}}{h}\right)$$

Also note that a preliminary test of gaussianity of the $\varepsilon'_{i,k}s$ can be made with the Jarque-Bera procedure which rejects the gaussianity at level α (for instance $\alpha = 5\%$) if :

$$n(n-1)\left(\frac{S^2}{6} + \frac{(K-3)^2}{24}\right) \geq \chi^2_{1-\alpha}(2) \tag{15}$$

where S and K are respectively the empirical skewness and hurtosis of the $\widehat{\varepsilon}_{i,k}$.

3.3 The CDR_i and their VaR

From proposition 4 we know that, in the Gaussian case, the conditional distribution of $Y_i^* = (Y'_{i,n+1-i}, ..., Y'_{i,N})'$ given $X_{i,n-i}$ and $d'X_{i,N} = 0$ is : $N(\mu_i^*, \Sigma_i^*)$ with :

$$\mu_i^* = m_i^* - \frac{c_i'^* a_i^*}{b_i^*}, \Sigma_i^* = \Omega_i^* - \frac{c_i^* c_i'^*}{b_i^*}$$

Denoting by T_i^* the Cholesky matrix satisfying $T_i^* T_i'^* = \Sigma_i^*$ we have :

$$Y_i^* = \mu_i^* + T_i^* \varepsilon_i^* \tag{16}$$

where the components $\varepsilon^*_{i,k}$ of ε_i^* follow independently $N(0,1)$. In the general case we can make the assumption that these components $\varepsilon^*_{i,k}$ follow independently a distribution estimated by the one obtained in section 3.2. At this stage it is important to stress the following property.

Proposition 5
For any distribution of the $\varepsilon^*_{i,k}$, the model $Y_i^* = \mu_i^* + T_i^*\varepsilon_i^*$ implies $X_{1,i,N} = X_{2,i,N}$.

Proof : Since the model $Y_i^* = \mu_i^* + T_i^*\varepsilon_i^*$ implies, for any distribution of the $\varepsilon^*_{i,k}$, the same first and second order moments of Y_i^* and X_i^* as in the Gaussian model, we have in particular $E(X_{1,i,N}-X_{2,i,N})=0$ and $V(X_{1,i,N}-X_{2,i,N})=0$ and therefore $X_{1,i,N}=X_{2,i,N}$ for any distribution of the $\varepsilon^*_{i,k}$. $\square$

Since the conditional expectation of Y_i^* given $X_{i,n-i}$ and $d'X_{i,N}=0$, remains equal to μ_i^*, the best prediction of $X_{n-i+k}, k\in\{1,\dots N-i+i\}$, remains $X_{i,n-i}+\sum_{j=1}^{k}\mu^*(i,n-i+j;\theta)$

The $CDR_i(n+1)$ is :

$$CDR_i(n+1)=X_{1,i,n-i}E_n[exp(f_i'Y_i^*)]-X_{1,i,n-i+1}E_{n+1}[exp(f_i'^*Y_i^{**})]$$

with :

$$Y_i^{**}=(Y'_{i,n-i+2},...,Y'_{i,N})'$$

or :

$$CDR_i(n+1)=X_{1,i,n-i}[E_nexp(f_i'Y_i^*)-exp(Y_{1,i,n-i+1})E_{n+1}exp(f_i'^*Y_i^{**})] \quad (17)$$

From (16) we get :

$$Y_{1,i,n-i+1}=\mu^*_{i,1}+T^*_{i,11}\varepsilon^*_{i,1}$$

and replacing the true value of θ_0 appearing in $\mu^*_{i,11}$ and $T^*_{i,11}$by $\widehat{\theta}_n$ we get :

$$Y_{1,i,n-i+1}=\widehat{\mu}^*_{i,1}+\widehat{T}^*_{i,11}\varepsilon^*_{i,1}$$

and :

$$\widehat{CDR}_i(n+1)=X_{i,n-i}[E_nexp(f_i'Y_i^*)-exp(\widehat{\mu}^*_{i,1}+\hat{T}^*_{i,11}\varepsilon^*_{i,1})E_{n+1}exp(f_i'^*Y_i^{**})]$$

If we do not take into account the estimation updating we can easily simulate $\widehat{CDR}_i(n+1)$ by simulating $\varepsilon^*_{i,1}$ in the distribution estimated in section 3.2 and by approximating both expectations by Monte Carlo using the values $\widehat{\theta}_n$ in the relevant components of μ_i^* and T_i^*.

If we want to take into account the estimation updating, for each simulation of $Y_{i,n+1-i}$ based on :

$$Y_{i,n-i+1}=\begin{pmatrix}\mu^*_{i,1}\\ \mu^*_{i,2}\end{pmatrix}+\begin{pmatrix}T^*_{i,11} & 0\\ T^*_{i,21} & T^*_{i,22}\end{pmatrix}\begin{pmatrix}\varepsilon^*_{i,1}\\ \varepsilon^*_{i,2}\end{pmatrix}$$

we must update $\widehat{\theta}_n$ into $\widehat{\theta}_{n+1}$ and, then compute the second expectation in (17) by Monte Carlo, replacing θ_0 by $\widehat{\theta}_{n+1}$ in the equations :

$$Y_i^{**} = \mu_i^{**} + T_i^{**}\varepsilon_i^{**}$$

The estimations of the $Var_i(\alpha)'s$ and the global $VaR(\alpha)$ are obtained from the empirical quantiles of M simulations of the $\widehat{CDR}_i(n+1)'s$ or :

$$\widehat{CDR}(n+1) = \sum_{i=1}^{n} \widehat{CDR}_i(n+1).$$

4 An application

We consider incurred claims and cumulated payments corresponding to a line of business Motor Body Liability-Insurance (the unit is 10^3 euros) [see appendix 3]. This line of business is highly volatile and therefore, not easy to model.

4.1 Estimation of the parameters

We begin with separate modellings of the rates of growth of cumulated payments and of incurred claims. For each variable we estimate by the nonlinear least square method a mean function, i.e. the corresponding component of $m(i, j, \theta)$, and a variance function i.e. the corresponding diagonal term of $\Omega(i, j, \theta)$. The mean function is assumed to be an affine function, with unknown coefficients (the components of θ) of basic functions of i and j, namely the identity function, the square function, the logarithmic function and the exponential. These mean functions are also assumed to be equal to zero if j is larger than a threshold J. The best set of basic functions and the optimal thresholds are selected according to the Akaike's criterion.

The basic functions retained, the estimation of their coefficients, the t-ratio statistics, and J are given in Tables 1 and 2. The variances, i.e. the diagonal terms of $\Omega(i, j)$, are assumed to be affine functions of the square of the corresponding mean. The estimation of the coefficients of this affine functions and the associated t-ratios are also given in Tables 1 and 2.

It is seen that all these estimations are highly significant. They will be used as starting values for the (pseudo) maximum likelihood method described above for the estimation of the CIP model. In this second stage the correlation function $\rho(i,j)$ appearing in $\Omega(i,j)$, i.e. the correlation between the two components of $Y_{i,j}$ (or $\xi_{i,j}$), has been taken into account; different specifications have been tested and a constant function has been retained. The estimations of the parameters and the corresponding t-ratio of the CIP model are also given in tables 1 and 2. It is interesting to see that these estimations are, in general, rather different from the initial values and this shows the importance of jointly taking into account the information contained in the cumulated payments and the incurred claims. It is also worth noting that all the coefficients are highly statistically significant. As mentioned in section 2.4, the CIP method also allows to propose an estimation for the ultimate development year N and we check that when some development profils are highly volatile like in the data considered here, the estimation of N may be large. In our case we find $N = 31$.

4.2 Values at the ultimate development year

In the CIP model the predicted values of the cumulated payments and of the incurred claims at the ultimate development date are, by construction, the same. It is interesting to compare these estimated ultimate values with the one provided by the Chain Ladder method applied to the cumulated

Table 1 : Cumulated payments ($J = 11$)

	Separate modeling			CIP modeling	
Mean	Function	Estimation	t-ratio	Estimation	t-ratio
	Intercept	-5.19	5.61	-3.51	4.59
	j	-0.24	3.34	-0.15	3.49
	$\log(j)$	-7.19	8.68	-5.25	5.97
	$\log(1+j)$	10.10	6.92	7.17	5.38
Variance	Intercept	0.01	1.15	0.009	6.22
	m^2	0.06	6.21	0.09	2.30

Table 2 : Incurred claims ($J = 4$)

	Separate modeling			CIP modeling	
Mean	Function	Estimation	t-ratio	Estimation	t-ratio
	Intercept	-0.23	3.14	-0.24	4.41
	i	0.08	4.72	0.06	5.97
	i^2	-0.005	4.09	-0.004	5.39
	j	0.24	3.84	0.24	4.82
	$\log(j)$	-0.63	4.76	-0.56	5.32
Variance	Intercept	0.004	2.16	0.002	6.79
	m^2	0.32	4.21	0.784	3.22
rho		0.28		0.26	3.41

payments, by the Chain Ladder method applied to the incurred claims and by the Munich-Re method. These values are displayed in table 3. The two Chain Ladder provide very different results, the total over the accident years being 73877.10^3 for the cumulated payments and 97954.10^3 for the incurred claims. The Munich Re method is similar to the Chain Ladder method for incurred claims. The CIP method provides, in general, values which are between the two chain ladders. In particular the total is 93471.10^3.

Table 3 : Values at the ultimate development year (10^3 euros)

	Chain Ladder Payment	Chain Ladder Incurred	Munich Re	CIP
1997	5909	7177	7177	7026
1998	3698	4711	4711	4588
1999	5688	9002	9038	8793
2000	5082	7040	7046	6960
2001	8803	9167	8908	9527
2002	6662	6976	6818	7146
2003	7344	11226	11281	10755
2004	5548	4427	4114	4565
2005	5842	4611	4123	4795
2006	7039	18770	20812	16839
2007	3441	3847	3631	3782
2008	3428	3980	3756	3600
2009	3365	4178	4070	3384
2010	2022	2836	2916	1704
Total	73877	97954	98405	93471

Figure 1 (resp.2) shows the predictions of the cumulated payments and of the incurred claims provided by the Chain Ladder and the CIP methods for accident year 2004 (resp 2008). In both cases the Chain Ladder method provides very different values for the two variables at the largest observed development horizon i.e. 14, and the ultimate common value proposed by the CIP method is between these two values.

Figure 1 : Prediction of the incurred and paid claims : year 2004.

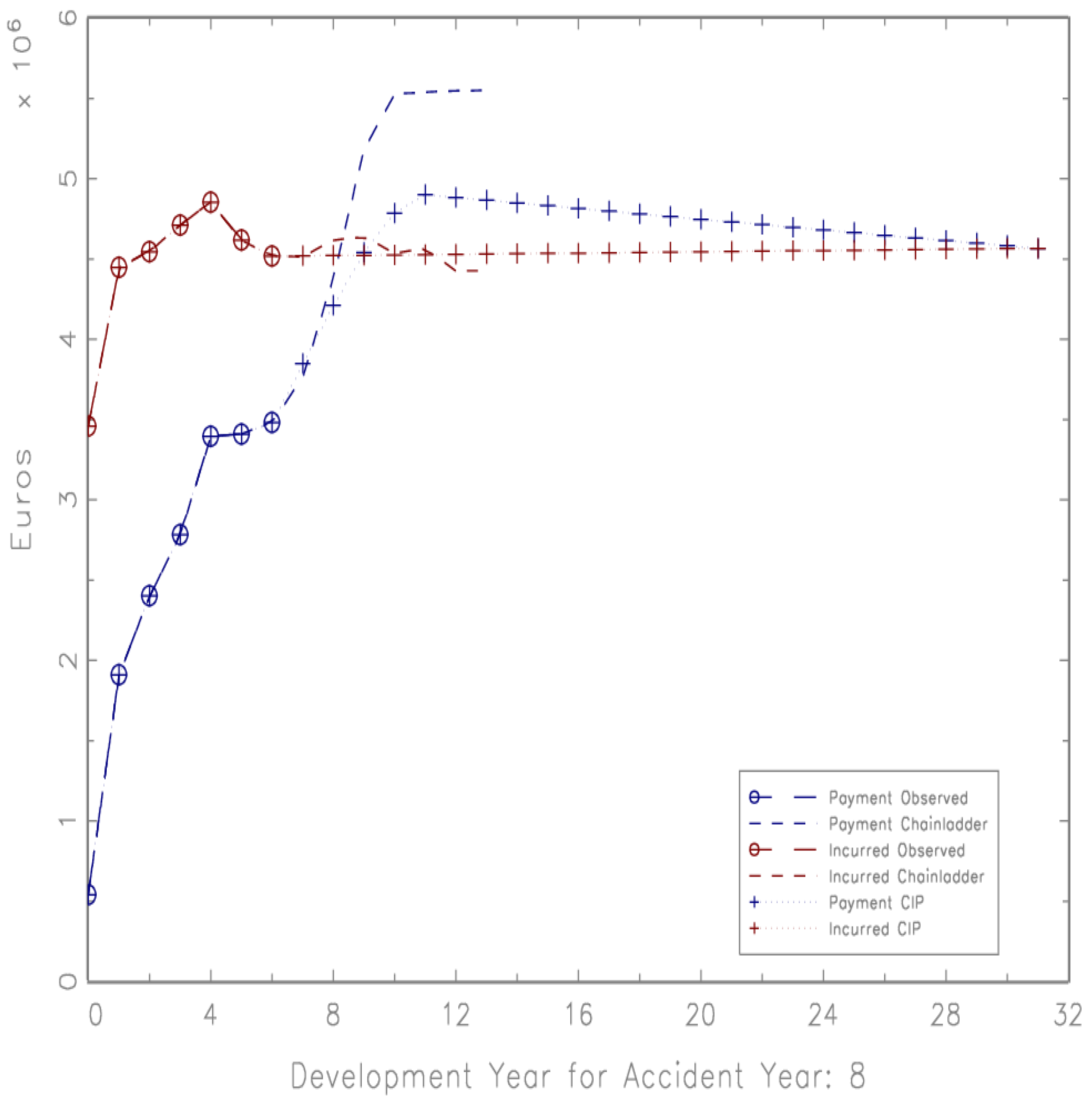
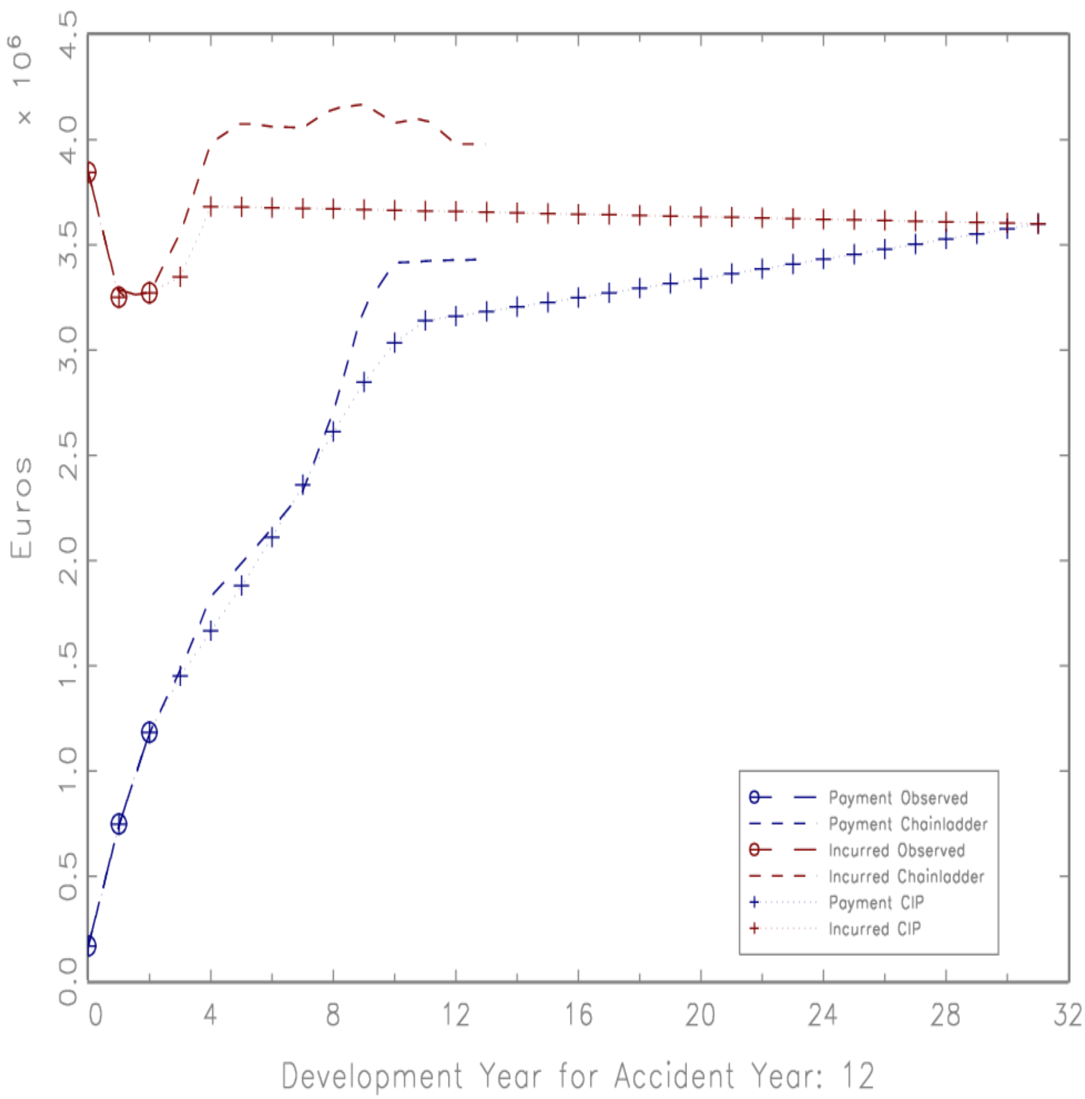


Figure 2 : Prediction of the incurred and paid claims : year 2008.



Figures 3 and 4 provides the whole prediction surfaces of the cumulated payments and of the incurred claims. By construction the profiles at the ultimate development horizon are identical.

Figure 3 : Prediction surface of the cumulated payments.

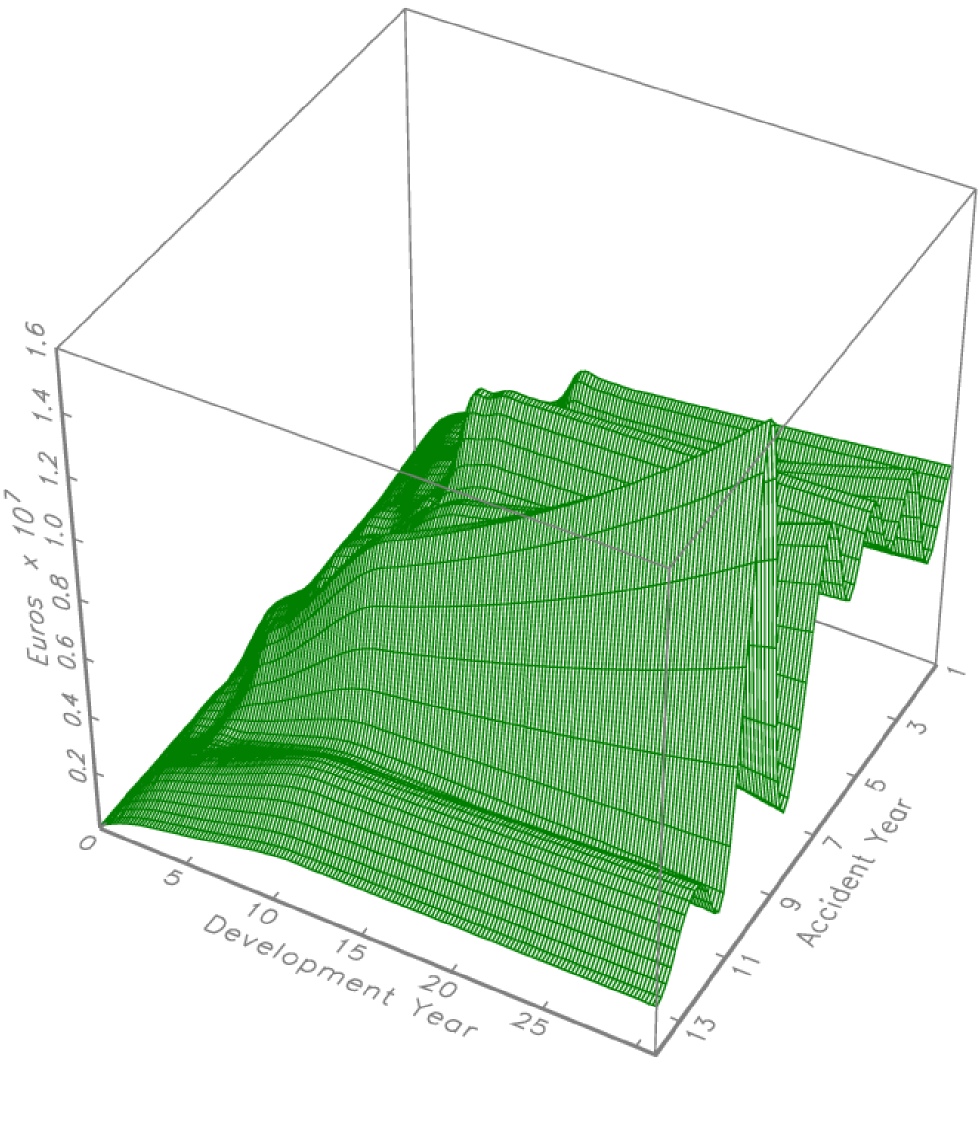
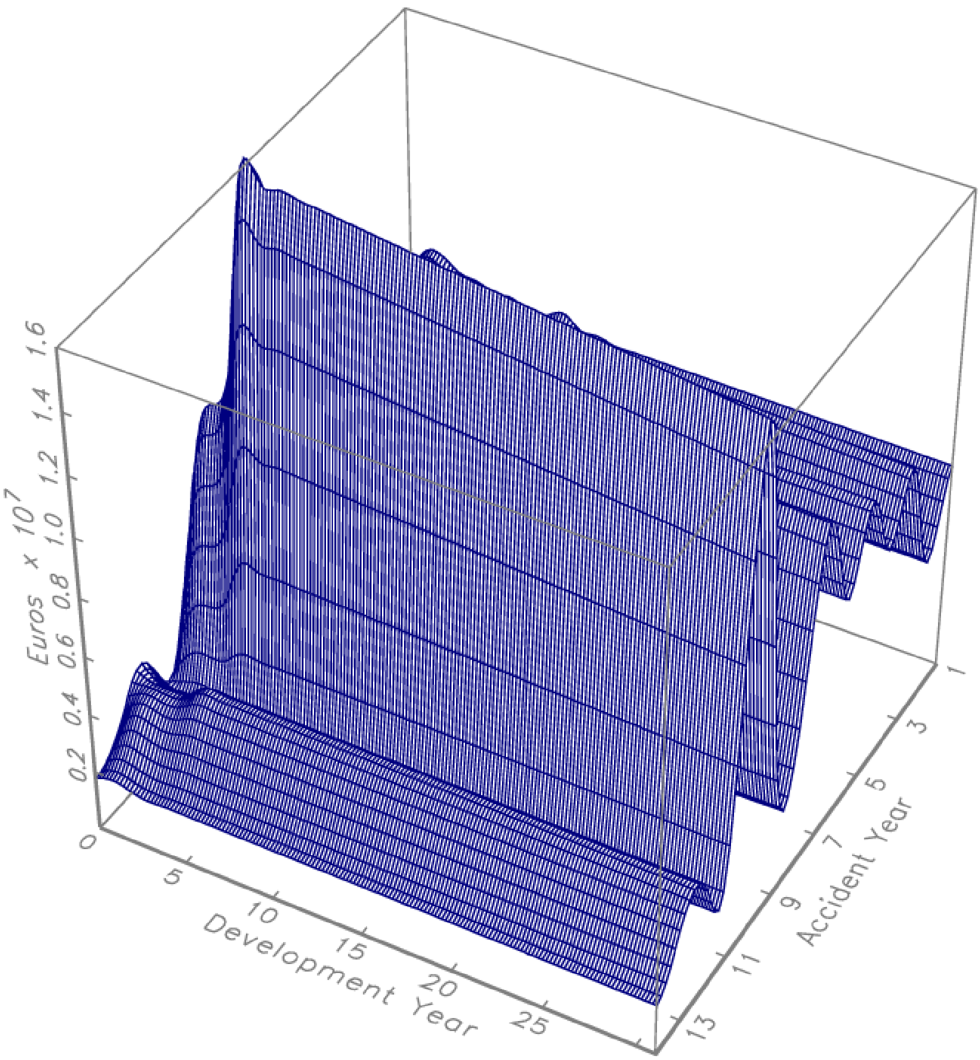


Figure 4 : Prediction surface of the incurred claims.



4.3 Values at risk of the CDR's

In a previous study only based on incurred claims [see Koenig, Le Moine, Monfort, Ratiarison (2015)] we have stressed the importance of two elements in the computation of the VaR's of CDR namely: the non-Gaussiarity of the distributions and the updating of the estimations. As we shall see, the importance of these features are strongly confirmed by the CIP method.

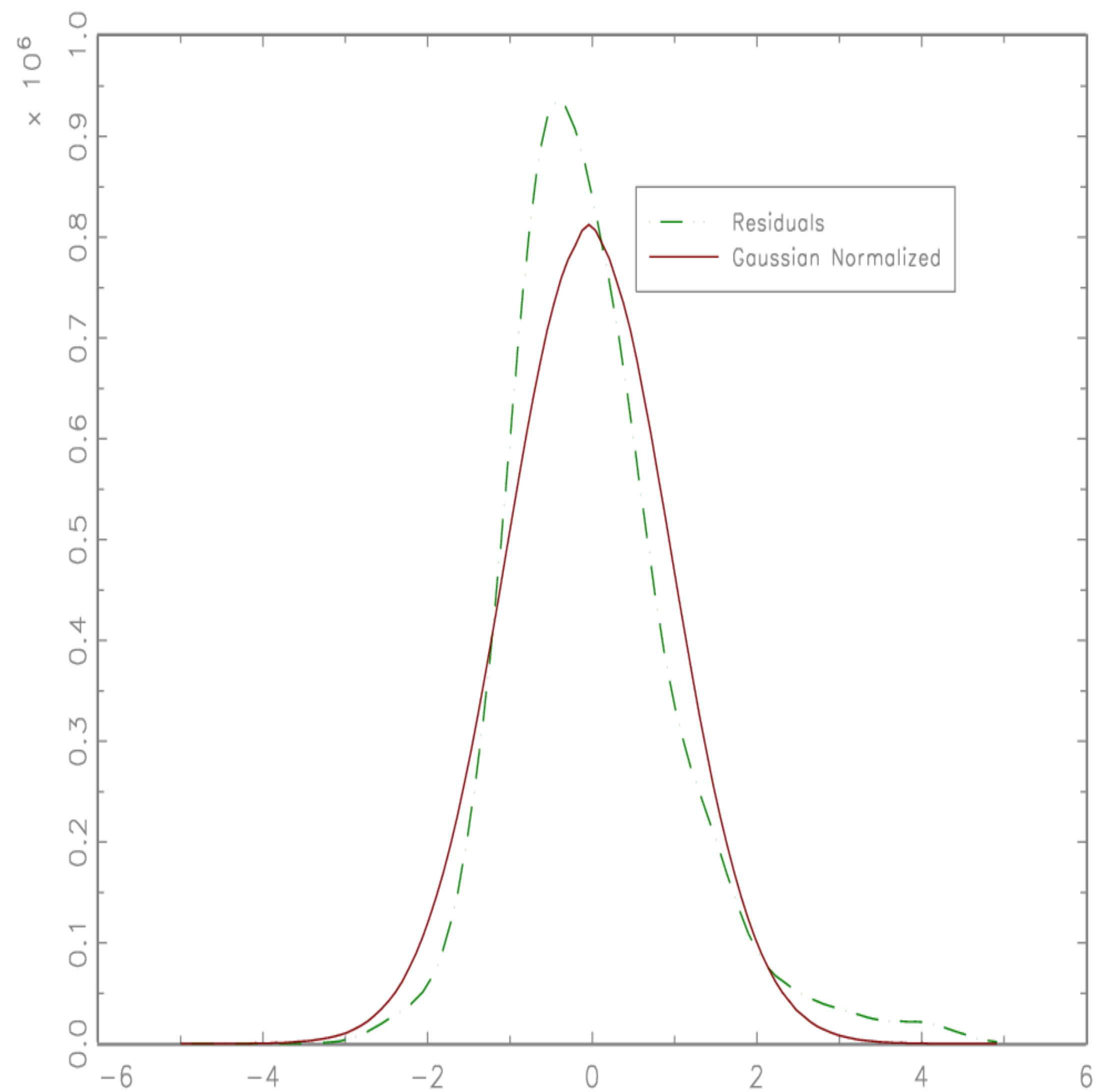
First let us test the normality of the components of the normalized vectors ε_i defined in equation (13) and estimated by $\hat{\varepsilon}_i$ defined in equation (14). Since $n = 14$ the Jarque-Bera statistic, given in (15), becomes :

$$91(\frac{S^2}{3} + \frac{(K-3)^2}{12})$$

where S and K are respectively the empirical skewness and kurtosis of the $\hat{\varepsilon}_{i,k}$.

If the errors are Gaussian the Jarque-Bera statistic is asymptotically distributed as $\chi^2(2)$ and the null hypothesis of normality should be rejected if the numerical value of this statistic is larger than the critical values, which are 4.6, 6.0, 9.2 for the 10%, 5% and 1% levels, respectively. Since the value found is 80.5 the normality assumption is very strongly rejected. This non-normality is confirmed by figure 5 showing the kernel based estimation of the density of the ε_i compared with the standard Gaussian density : a much ticker right tail is observed.

Figure 5 : Estimated density function of the residuals.



If follows that the appropriate computation of the VaR's of the CDR's should not assume normality and therefore should be based on the method described in section 3.3. Moreover it is important to measure the impact of the updating of the estimations of the parameters. Table 4 gives the results.

Table 4 : VaR's of the CDR's (10^3 euros)

	without updating		with updating	
	Gaussian	Non Gaussian	Gaussian	Non Gaussian
1997	914	1252	901	1281
1998	583	827	555	781
1999	1118	1610	1156	1606
2000	895	1277	916	1181
2001	1251	1751	1222	1669
2002	919	1305	942	1352
2003	1409	1929	1457	2096
2004	581	848	611	866
2005	619	894	649	871
2006	2201	3108	2451	3316
2007	1459	2187	1263	1784
2008	531	765	604	885
2009	490	709	749	1105
2010	419	612	668	1001
Sum of VaR's	13394	19081	14150	19803
Global VaR	**3779**	**4954**	**5031**	**6695**

Let us consider the global 99.5% VaR. Wrongly assuming normality leads to a VaR equal to 3779.10^3 instead of 4954.10^3 when there is no updating and a VaR equal to 5031.10^3 instead of 6695.10^3 when there is updating. The price to pay for wrongly assuming normality is very high : an under-estimation of approximately 25% .

The price to pay for omitting updating is of the same order of magnitude. It moves from 3779.10^3 to 5031.10^3 in the Gaussian case and from 4954.10^3 to 6695.10^3 in the non Gaussian case. Cumulating both mistakes lead to an under-estimation of approximately 44%.

5 CONCLUDING REMARKS

We proposed a flexible statistical modelling, called the CIP method, allowing to take into account simultaneously the payments and the incurred claims in the prediction of future claims. This method is semi-parametric since it does not assume a precise shape of the distributions but only concentrates on the first two moments. In particular normality of the growth rates, i.e. log-normality of the levels, is not assumed and is in fact strongly rejected in our application. Moreover our CIP method also allows to estimate the ultimate development year, the CDR's (Claim Development Results) and their VaR's (Value at Risk) which are measures of reserve risk recommended by the regulatory authorities. The techniques derived in this paper could be extended in several directions. In particular it would be interesting to derive a CIP method treating simultaneously several business lines. This kind of development is left for future research.

APPENDIX 1
Proof of proposition 1

Let us first consider the joint distribution of $\begin{pmatrix} \tilde{Y}_i \\ d'X_{i,N} \end{pmatrix}$. Since $X_{i,N} = X_{i,0} + \sum_{j=1}^{N} Y_{i,j}$, this joint distribution (given $X_{i,0}$) is Gaussian.

Its mean is :

$$\begin{pmatrix} \tilde{m}_i \\ d'X_{i,0} + d' \sum_{j=1}^{N} m_{i,j} \end{pmatrix} = \begin{pmatrix} \tilde{m}_i \\ a_i \end{pmatrix}$$

and its variance-covariance matrix is :

$$\begin{pmatrix} \Omega_{i,1} \ldots & 0 & \tilde{c}_i \\ 0 & \Omega_{i,N} & \\ \tilde{c}_i' & & b_i \end{pmatrix} = \begin{pmatrix} \tilde{\Omega}_i & \tilde{c}_i \\ \tilde{c}_i & b_i \end{pmatrix}$$

with :

$$\begin{aligned} b_i &= V(d'X_{i,N}) = \sum_{j=1}^{N} d'\Omega_{i,j} d \\ \tilde{c}_i &= cov(\tilde{Y}_i, d'X_{i,N}) \\ &= \begin{pmatrix} c_{i,1} \\ \vdots \\ c_{i,N} \end{pmatrix} \end{aligned}$$

and $c_{ij} = \Omega_{i,j} d$ and therefore $\tilde{c}_i = \tilde{\Omega}_i F_N d$, with $F_N = (I_2, ..., I_2)'$.

Applying a standard formula for conditional Gaussian distributions we see that the conditional distribution of $\tilde{Y}_i$ given $d'X_{i,N} = 0$ is :

$$N\left(\tilde{m}_i - \frac{\tilde{c}_i a_i}{b_i}, \tilde{\Omega}_i - \frac{\tilde{c}_i \tilde{c}_i}{b_i} \right). \square$$

APPENDIX 2
Proof of proposition 3

Lemma

Let β a vector such that $\parallel \beta \parallel \neq 1$ (with $\parallel \beta \parallel^2 = \beta'\beta$), the matrix $I - \beta\beta'$ is invertible and

$$(I - \beta\beta')^{-1} = I + \frac{\beta\beta'}{1- \parallel \beta \parallel^2}.$$

Proof :

$$\begin{aligned}(I - \beta\beta')\left(I + \frac{\beta\beta'}{1- \parallel \beta \parallel^2} \right) &= I - \beta\beta' + \frac{\beta\beta'}{1- \parallel \beta \parallel^2} - \frac{\beta\beta' \parallel \beta \parallel^2}{1- \parallel \beta \parallel^2} \\ &= I \qquad \square\end{aligned}$$

Let us now consider the matrix :

$$\Omega_i - \frac{c_i c_i'}{b_i} = \Omega_i^{1/2} \left(I - \frac{\Omega_i^{-1/2} c_i}{b_i^{1/2}} \frac{c_i' \Omega_i^{-1/2}}{b_i^{1/2}} \right) \Omega_i^{1/2}$$

setting $\beta_i = \dfrac{\Omega_i^{-1/2} c_i}{b_i^{1/2}}$ we get :

$$\Omega_i - \frac{c_i c_i'}{b_i} = \Omega_i^{1/2} (1 - \beta_i \beta_i') \Omega^{1/2}$$

and applying the lemma we get :

$$\begin{aligned}\left(\Omega_i - \frac{c_i c_i}{b_i} \right)^{-1} &= \Omega_i^{-1/2} \left(I + \frac{\beta_i \beta_i'}{1- \parallel \beta_i \parallel^2} \right) \Omega_i^{-1/2} \\ &= \Omega_i^{-1} + \frac{\Omega_i^{-1} c_i c_i' \Omega_i^{-1}}{b_i \left(1 - \dfrac{c_i' \Omega_i^{-1} c_i}{b_i} \right)} \\ &= \Omega_i^{-1} + \frac{\Omega_i^{-1} c_i c_i' \Omega_i^{-1}}{b_i - c_i' \Omega_i^{-1} c_i} \qquad \square\end{aligned}$$

APPENDIX 3

Payments (10^3 euros)

455	1941	2647	3006	3248	3608	3855	4215	4427	4811	5890	5910	5907	5910
659	1780	2410	2954	3103	3192	3302	3596	3670	3678	3678	3680	3697	
480	1536	2340	2831	3343	3613	3920	4815	5499	5643	5664	5678		
183	1433	2039	2449	2985	3195	3637	3795	4597	4913	5061			
131	1372	2081	3001	3289	3537	4105	4211	4874	8227				
139	1266	1975	2385	2647	3177	3508	3691	5267					
372	1371	2067	2614	3869	4467	4774	4984						
539	1909	2401	2781	3394	3409	3482							
212	1042	1889	2396	3204	3381								
203	1098	1615	2184	3751									
209	804	991	1486										
168	747	1183											
161	801												
110													

Incurred claims (10^3 euros)

7430	8289	8223	7849	7786	7561	7623	7284	7400	7696	7317	7318	7177	7177
6654	7319	6453	6340	6318	6014	5375	5414	5385	5078	5013	4933	4711	
6555	7795	7394	7990	7540	8596	8314	8713	8931	8995	9072	9277		
5804	6573	6248	6187	6462	6308	7238	7104	7301	7481	7213			
5527	6644	7551	9198	9008	9423	9732	9324	9561	9601				
4066	5945	6537	6332	6272	6797	6651	7077	7272					
3301	5169	5808	8083	11294	11922	11490	11441						
3459	4446	4544	4709	4853	4617	4517							
2376	4073	4065	4337	4817	4728								
6035	10398	11191	13269	18773									
2470	3260	3452	3439										
3844	3251	3271											
2733	3362												
1797													

References

Boor J. (2006), "Estimating tail development factors : what to do when the triangle runs out", CAS Forum Winter 2006, 345-390.

Choirat C. and Seri R. (2012), "Discrete Parameter Models, Statistical Science", 27, 278-293.

Gourieroux C., Monfort A. and Trognon., (1984), "Pseudo Maximum Likelihood Methods : Theory", Econometrica, 52, 680-700.

Haliwell I.J. (1997), "Conjoint prediction of paid and incurred losses, CAS Forum Summer".

Haliwell I.J. (2009), "Modeling paid and incurred losses together, CAS Forum Spring".

Happ S., Merz M., Wuthrich M.V. (2011), "Claims development result in the paid incurred claim reserving method". Preprint.

Happ S., Wuthrich M.V. (2011), "Paid-incurred chain reserving with dependence". Preprint.

Koenig E., Le Moine P., Monfort A. and Ratiarison E.,, (2015) "Evaluating reserve risk in a regulatory perspective," Journal of Insurance Issues, 38(2), 151-183.

Merz M. and Wuthrich M.V. (2010-a), "Paid-Incurred chain claims reserving method", Insurance Mathematics and Economics, 46, 568-579.

Merz M. and Wuthrich M.V. (2010-b), "Estimation of tail development factors in the Paid-Incurred chain reserving method". Preprint.

Posthuma B., Cator E.A., Veerkamp W., Van Zwet E.W. (2008), "Combined analysis of paid and incurred losses", Preprint.

Quang C., Mack T., (2004), "Munich chain ladder", Blatter DGVFM Band XXVI, 597-630.

Venter G.G., (2008), "Distribution and value of reserves using paid and incurred triangle", CAS Forum Fall.

Wuthrich M.V., Merz M. (2008), "Stochastic claims reserving methods in insurance", Wiley.