Dynamic Claims Reserving in Non-Life Insurance: A State Space Approach with Kalman Filtering and Monte Carlo Forecasting

By

Dr. Ahmed Samy Said El Azab

Department of Insurance and Actuarial Science, Faculty of Commerce, Cairo University

ahmed_samy@foc.cu.edu.eg

Dr. Raghda Ali Abdelrahman

Department of Insurance and Actuarial Science, Faculty of Commerce, Cairo University

raghda.ali.abd@gmail.com

Dr. Zahra Salah Eldin

Department of Insurance and Actuarial Science, Faculty of Commerce, Cairo University

z.saladin@foc.cu.edu.eg

Scientific Journal for Financial and Commercial Studies and Research (SJFCSR)

Faculty of Commerce – Damietta University

Vol.6, No.2, Part 1., July 2025

APA Citation

El Azab, A. S. S.; **Abdelrahman**, R. A. and **Eldin,** Z. S. (2025). Dynamic Claims Reserving in Non-Life Insurance: A State Space Approach with Kalman Filtering and Monte Carlo Forecasting, ***Scientific Journal for Financial and Commercial Studies and Research***, Faculty of Commerce, Damietta University, 6(2)1, 1055-1085.

Website: https://cfdj.journals.ekb.eg/

Dynamic Claims Reserving in Non-Life Insurance: A State Space Approach with Kalman Filtering and Monte Carlo Forecasting

Dr. Ahmed Samy Said El Azab; Dr. Raghda Ali Abdelrahman and Dr. Zahra Salah Eldin

Abstract

Accurate estimation of claims reserves is essential to maintaining the financial stability of non-life insurers, particularly in markets subject to structural volatility and evolving regulatory standards. Traditional stochastic reserving models such as the Mack model are widely used for their transparency and analytical simplicity, yet they often fall short in environments characterized by irregular reporting, dynamic settlement patterns, and macroeconomic disruption.

To overcome these challenges, this research proposes a Scalar State Space Model (SSM) integrated with Kalman filtering. Applied to cumulative paid claims data from an Egyptian motor insurance portfolio, the model captures latent calendar-year effects and enables recursive reserve updating as new data becomes available. Through Monte Carlo simulation, the SSM produces full predictive distributions of future liabilities, offering a comprehensive view of reserve uncertainty.

Comparative analysis against the benchmark Mack model shows that the SSM delivers more stable reserve estimates and better reflects underlying risk, especially in recent accident years where uncertainty is most pronounced. Unlike the Mack model's reliance on fixed development patterns and independence assumptions, the SSM dynamically models time-varying processes and structural shocks.

The results highlight the SSM's advantages in adaptability, robustness, and regulatory alignment, making it a compelling alternative for insurers operating under IFRS 17 and similar solvency-focused regimes. These findings advocate for the integration of dynamic stochastic reserving methods into the actuarial toolkit, particularly in emerging markets facing data volatility and structural transformation.

Keywords: State Space Model; Kalman Filter; IBNR; Mack Model; Monte Carlo Simulation; Forecast Accuracy; Stochastic Reserving.

1. Introduction

Claim reserving is a central pillar in non-life insurance operations, particularly for estimating liabilities associated with claims that have been Incurred But are not yet fully Reported or settled (IBNR). Accurate reserving is essential not only for meeting policyholder obligations, but also for maintaining solvency, ensuring regulatory compliance, and supporting sound pricing and reinsurance strategies.

Over the past decades, actuaries have developed a wide array of reserving techniques based on historical claims development data, typically structured in run-off triangles. These techniques are broadly categorized into *deterministic* and *stochastic* methods. Deterministic approaches, such as the widely used Chain Ladder method, are simple and interpretable but are often criticized for their sensitivity to outliers and lack of explicit measures of uncertainty (England & Verrall, 2002). As a response, stochastic reserving models have gained popularity, offering a more rigorous framework by incorporating randomness in the claims process (Mack, 1993); (Wüthrich & Merz, 2008).

Among the early stochastic models, the Mack model (Mack, 1993) and (Mack, 1999) stands out for providing distribution-free estimates of reserve variability while maintaining consistency with the Chain Ladder's point estimates. Later advancements introduced Generalized Linear Models (GLMs), allowing actuaries to flexibly model incremental claims while accounting for heteroscedasticity and overdispersion see (England & Verrall, 2002). Despite their strengths, both Chain Ladder and GLM-based approaches are essentially static, assuming fixed development patterns over time.

However, insurance claims development often exhibits time-dependent behavior, affected by evolving business practices, macroeconomic conditions, calendar year trends, and legal reforms. To address such complexities, researchers have turned to dynamic stochastic models, most notably State Space Models (SSMs). These models, initially introduced to actuarial reserving by (DeJong & Zehnwirth, 1983) and later extended by (Verrall, 1994) and (Taylor, et al., 2003), represent claims development as a latent process evolving over time. SSMs are estimated using Kalman filtering and smoothing algorithms, which enable real-time updating, better forecasting accuracy, and full predictive distributions.

The potential of such advanced models is particularly relevant in the context of Egypt's insurance industry, which has been undergoing significant transformation. The Financial Regulatory Authority (FRA) has intensified efforts to modernize actuarial practices, improve risk management, and align with global reporting standards like IFRS 17. This new regime emphasizes a market-consistent, forward-looking, and uncertainty-aware approach to reserving; one that traditional deterministic models struggle to fulfill. Additionally, the Egyptian market faces data limitations, especially for long-tail lines of business, which makes the flexibility and robustness of dynamic models like SSMs even more valuable.

This research contributes to the ongoing development of reserving practices by applying a scalar State Space Model with Kalman filtering to a cumulative paid claims triangle. The model's performance is empirically assessed and compared with the benchmark Mack model, focusing on forecast accuracy, reserve uncertainty, and practical implementation aspects. The analysis provides insights for Egyptian insurers and regulators into the viability of adopting dynamic reserving techniques that are both statistically robust and compliant with evolving global standards.

2. Research Problem

Traditional claims reserving techniques, most notably the deterministic Chain Ladder method and its stochastic extension, the Mack model; have long been pillars of actuarial practice due to their simplicity and minimal data requirements. However, these methods rely on strong assumptions: consistent development factors across time, independence between accident years, and homogeneous claim settlement behavior. In increasingly dynamic insurance environments, particularly in emerging markets like Egypt, these assumptions are often violated.

The Egyptian non-life insurance market is characterized by irregular reporting patterns, volatile claim amounts, and systemic disruptions, such as inflation surges, legal reforms, and operational shifts. These complexities are especially prominent in high-volume lines like motor and medical insurance, where claim settlement processes can be highly nonlinear and reactive. Consequently, models that assume stable development patterns fail to accommodate the true nature of claim behavior, leading to unreliable reserve estimates and heightened financial risk.

This issue is further amplified by the country's impending adoption of IFRS 17. Unlike previous frameworks, IFRS 17 demands that insurers incorporate forward-looking cash flows, explicit risk margins, and discounting mechanisms in reserve estimation. Classical methods, which focus on point estimates without accounting for uncertainty or process dynamics, are fundamentally misaligned with these requirements.

The limitations of static reserving models become evident when examining real-world claims data, as illustrated in ***Figure 2.1.*** This heatmap shows cumulative paid claims across accident years (rows) and development years (columns) for motor insurance. In a stable claims development process, we would expect smooth, horizontally consistent gradients, reflecting uniform claim settlement behavior over time. However, the heatmap reveals pronounced irregularities, including abrupt shifts in intensity, plateaus, and nonlinear growth patterns, particularly in recent accident years.

These disruptions reflect structural inconsistencies in claim reporting and settlement dynamics, potentially due to changes in operational practices, inflationary effects, or economic shocks. Such volatility violates key assumptions of traditional deterministic models like the Chain Ladder, which rely on stable, independent development factors across accident years. Consequently, applying these models uncritically risks significant misestimation, either over- or under-reserving.

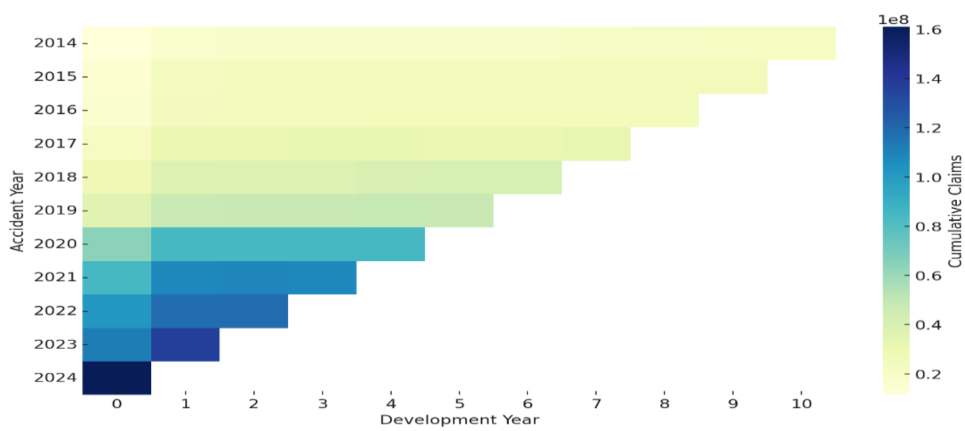

Figure 2.1. Volatility in Cumulative Claims Developments

These empirical observations highlight a critical need for more adaptive, probabilistic models that can handle uncertainty, evolving trends, and unobservable processes. State Space Models (SSMs); particularly when combined with Kalman filtering, offer a compelling alternative. They allow for the decomposition of claim development into latent state processes and observed noisy data, enabling real-time updates and robust forecasting.

This research addresses the existing gap by empirically evaluating a scalar SSM against the benchmark Mack model using real Egyptian claims data. By comparing the two approaches in terms of forecast accuracy, reserve volatility, and IFRS 17 compatibility, the research aims to determine whether SSMs offer a more resilient and regulation-ready framework for reserving in volatile and data-limited markets.

3. Research Objectives

The primary objective of this research is to evaluate and compare the effectiveness of a scalar SSM with Kalman filtering against the classical Mack model for IBNR claims reserves in non-life insurance. The focus is on assessing forecasting accuracy and uncertainty quantification using Monte Carlo simulation.

This research is guided by the following specific objectives:

1- **To identify the limitations of the Mack model** in capturing volatile and irregular claims development patterns in the Egyptian motor insurance market.

2- **To construct and estimate a scalar State Space Model** formulated on the cumulative paid claims' triangle, incorporating latent calendar-year effects that evolve stochastically over time.

3- **To apply the Kalman filter and smoother** for real-time estimation and dynamic updating of latent development trends, enabling recursive forecasting of future claims.

4- **To estimate IBNR reserves and predictive uncertainty** under the SSM using Monte Carlo simulation, thereby obtaining a full empirical distribution of future liabilities rather than point estimates alone.

5- **To compare the performance of the Mack model and the SSM** based on:
   - Point estimate of IBNR reserves,
   - Forecasting error as measured by Mean Squared Error of Predication (**MSEP**),
   - Coefficient of variation as a proxy for reserve volatility,
   - Ability to capture tail risk through predictive distributions.

6- **To evaluate the practical implications of dynamic modeling** for insurers operating in volatile or structurally shifting environments, with particular emphasis on the benefits of probabilistic and adaptive reserving approaches.

4. Literature Review

Over the past two decades, actuarial reserving has undergone a notable transformation, shifting from deterministic approaches like the Chain Ladder method to more flexible, stochastic, and dynamic models. Among these, State Space Models (SSMs) have gained substantial attention for their ability to capture latent structures, evolving claim behaviors, and data irregularities.

Early contributions, such as those by (Alpuim & Ribeiro, 2003), introduced state space models as a dynamic alternative to static triangle-based methods. Their approach allowed for the decomposition of observed claims data into latent and observable components, offering improved estimation of reserves and uncertainty. Similarly, (Pang & He, 2012) and (Atherino, et al., 2010) expanded this framework by emphasizing recursive estimation and restructured triangle data to accommodate time-series-based modeling of IBNR reserves.

A significant leap in methodological development came from (Chukhrova & Johannssen, 2017), who applied Kalman filtering and smoothing techniques within an SSM context. Their work provided theoretical and empirical insights into how SSMs can outperform static models like the Mack model, particularly by enabling real-time updating of reserves and by producing full predictive distributions. This was further reinforced in their later review (Chukhrova & Johannssen, 2021), which systematically classified Gaussian and non-Gaussian SSMs and highlighted their capacity to handle both observation and process noise.

In parallel, (Peters, et al., 2017) advanced the Bayesian formulation of SSMs, emphasizing their use in generating full predictive distributions for reserves. They demonstrated how such models accommodate structural changes in the data while maintaining transparency and interpretability—key demands under modern solvency and reporting regimes.

Building on these foundations, (Hendrych & Cipra, 2021) provided a practical framework for real-time reserving using Kalman filtering, focusing on model diagnostics and empirical validation. Their findings affirmed that SSMs offer superior adaptability and robustness, especially in the presence of structural volatility or data irregularities. (Nomura & Matsumori, 2024) introduced dynamic factor models within this paradigm, offering dimensionality reduction and more interpretable reserve forecasts in high-dimensional settings.

The emergence of machine learning and granular reserving methods has also influenced recent literature. (Taylor, 2019) contrasted traditional models with individual claim-level approaches, emphasizing the potential of hybrid techniques that integrate machine learning with classical actuarial methods. Similarly, (chwab & Schneider, 2024) proposed a novel neural network architecture to enhance the prediction of loss amounts incurred for reported but not settled (RBNS) claims, demonstrating improved accuracy over standard benchmark models like the chain ladder approach. In parallel, (Selukar, 2025) showcased how tools like SAS can operationalize SSMs with Kalman filtering for reserve estimation in real insurance portfolios.

Despite this growing body of work, most applications remain focused on mature insurance markets with relatively stable data environments. As noted by (Gogol, 2019), emerging markets face unique challenges such as volatile claims development, sparse data, and shifting regulations, all of which may invalidate the assumptions of classical models like Mack or Chain Ladder. However, empirical studies applying SSMs in these settings, especially Egypt, are still scarce.

Although the literature has extensively highlighted the theoretical and practical strengths of State SSMs integrated with Kalman filtering in stochastic claims reserving, empirical validation of their application in emerging markets, such as Egypt, remains scarce. These markets present significant challenges, including high data volatility, irregular reporting, and evolving regulatory requirements such as IFRS 17, which emphasizes probabilistic reserving approaches with specified confidence levels. In the Egyptian insurance landscape, particularly within high-frequency lines like motor, marine, and medical insurance, claims development is often unpredictable, undermining the foundational assumptions of conventional models. Nevertheless, there has been limited empirical research assessing the real-world performance of SSMs in comparison to established models like the Mack model within this volatile and transitioning environment.

5. Research Methodology

To address the evolving, uncertain, and latent nature of claims development, this research adopts a scalar State Space Model (SSM) framework inspired by (Chukhrova & Johannssen, 2017) and (Selukar, 2025). The model distinguishes between observed cumulative claims and unobserved latent states, such as development trends and calendar-year effects, enabling a dynamic view of the reserving process. Unlike traditional static methods, the SSM explicitly accounts for both **state (processed) noise** and **observation (measurement) noise**, making it well-suited for volatile environments with reporting irregularities. This approach allows for sequential estimation and real-time updating of reserves using Kalman filtering and smoothing, offering improved adaptability and accuracy in forecasting future claims.

5.1. Model Structure

We consider cumulative claims $\boldsymbol{C_{i,j}}$ where $\boldsymbol{i}$ is the accident year and $\boldsymbol{j}$ is the development year. The calendar year is defined by $\boldsymbol{t = i + j}$, and the scalar state space model is formulated in calendar-year format to reflect structural influences that affect entire diagonals of the run-off triangle.

5.1.1. State Equation

In state space modeling, the **state equation** plays a critical role in capturing the evolution of unobservable processes that drive the observable claim development. Specifically, in the context of claims reserving, these latent processes may reflect structural calendar year effects such as inflation, regulatory changes, or shifts in claim settlement practices—factors that are not directly observable in the claims triangle but influence its progression over time.

The state equation provides a mathematical framework for modeling how these latent components evolve across calendar years. It defines how the unobserved state at time $\boldsymbol{t}$, denoted $\boldsymbol{\theta_t}$, depends on its value in the previous year $\boldsymbol{\theta_{t-1}}$, subject to random fluctuations or shocks. This allows the model to adapt flexibly to gradual or abrupt changes in the development environment.

The evolution of the latent state is described by the **state equation**:

$$\boldsymbol{\theta_t = G_t\, \theta_{t-1} + \omega_t\,, \quad \omega_t \sim N(0, Q)} \qquad (1)$$

where:

- $\boldsymbol{\theta_t} \in \mathbb{R}$ is the scalar latent state at calendar year $\boldsymbol{t}$,
- $\boldsymbol{G_t} \in \mathbb{R}$ is the system coefficient governing the persistence of the state,
- $\boldsymbol{\omega_t}$ is the system noise (state innovation), modeled as a white noise process with zero mean and constant variance $\boldsymbol{Q}$,
- $\boldsymbol{Q > 0}$ is the state noise variance, capturing the uncertainty or variability in the evolution of the latent state.

In this research, we adopt the **random walk** specification, which assumes that the state follows a non-stationary path with perfect persistence. This is achieved by setting:

$$\boldsymbol{G_t = 1\; for\; all\; t}$$

leading to the simplified form:

$$\boldsymbol{\theta_t = \theta_{t-1} + \omega_t} \qquad (2)$$

The random walk assumption implies that the latent process evolves as the cumulative sum of independent shocks over time. It reflects a system where structural changes (e.g., gradual inflationary effects or regulatory drift) are not mean-reverting, making this formulation particularly suitable for long-term forecasting in claims reserving. This allows the model to track evolving trends without forcing the latent state to return to a fixed long-term average.

The estimated sequence $\{\boldsymbol{\theta_t}\}$ thus serves as a smooth, calendar-year-driven representation of the underlying claim development trend, which is then linked to observed claims data through the observation equation described in the next section.

5.1.2. Observation Equation:

While the state equation governs the latent calendar-year trend in claim development, the **observation equation** links this latent process to the actual data observed in the claims triangle—typically cumulative claims amounts.

In the scalar state space formulation, we assume that each observed log-cumulative claim value $\boldsymbol{y_t = \log (C_{i,j})}$ at **cell**$(\boldsymbol{i,j})$, where $\boldsymbol{t = i + j}$ (calendar year), is generated as a noisy measurement of the latent state $\boldsymbol{\theta_t}$:

$$\boldsymbol{y_t = \theta_t + v_t}\,, \qquad \boldsymbol{v_t \sim N\,(0, R)} \tag{3}$$

where:

- $\boldsymbol{y_t} \in \mathbb{R}$ is the observed log-transformed cumulative claim at development cell$(\boldsymbol{i,j})$,
- $\boldsymbol{v_t \sim N\,(0, R)}$ is the observation noise (i.e., deviation of actual claims from the latent trend),
- $\boldsymbol{R > 0}$ is the observation variance, capturing volatility in the observed claims not explained by the calendar year effect.

In practice, each calendar year $\boldsymbol{t}$ corresponds to a diagonal in the claims triangle and may contain multiple observed cell$(\boldsymbol{i,j})$ where $\boldsymbol{i + j = t}$. In such cases, the equation generalizes to:

$$\boldsymbol{y_t = 1_{nt}\, \theta_t + v_t}\,, \quad \boldsymbol{v_t \sim N\,(0, RI_{nt})} \tag{4}$$

Here:

- $\boldsymbol{y_t} \in \mathbb{R}^{n_t}$ is a vector of observed log claims for calendar year $\boldsymbol{t}$,
- $\boldsymbol{1_{nt}}$ is a column vector of ones,
- $\boldsymbol{v_t}$ is a vector of independent observation errors.

This structure assumes all claims in the same calendar year are influenced by a common latent driver $\boldsymbol{\theta_t}$, justifying the use of a shared state for each diagonal.

5.2. Model Assumptions

The scalar SSM used in this research is built upon a set of assumptions that ensure both theoretical soundness and computational tractability. These assumptions are crucial for the effective application of the Kalman Filter in estimating the latent development process and forecasting future claims liabilities.

5.2.1. Linearity of System and Observation Relationships

The model assumes that the underlying dynamics of the claims process, as well as the link between the latent states and the observed cumulative claims, are linear. This linearity allows for efficient state estimation using recursive filtering techniques. It also reflects a practical simplification that aligns with many existing actuarial reserving models, such as Chain Ladder and Mack models.

5.2.2. Normally Distributed Innovations

The disturbances affecting both the system dynamics and the measurement process are assumed to be Gaussian:

- The **state innovation** $\boldsymbol{\omega_t \sim N(0, Q)}$, representing unobserved changes in the latent process over time.
- The **observation noise** $\boldsymbol{v_t \sim N\,(0, R)}$, capturing deviations of the observed claims from the underlying latent structure.

These Gaussian assumptions facilitate analytical derivation of the Kalman Filter updates and enable closed-form expressions for forecast error variances.

5.2.3. Time-Invariant Variance Parameters

The variances $\boldsymbol{Q}$ and $\boldsymbol{R}$ are assumed to be constant across all time periods. This implies homoscedasticity of both the system and observation noise, making the model stable and interpretable. Extensions to time-varying variance structures can be considered for modeling volatility dynamics but are beyond the scope of the current research.

5.2.4. Mutual Independence of Noise Terms

It is assumed that the process noise $\boldsymbol{\omega_t}$ and observation noise $\boldsymbol{v_t}$ are mutually independent and uncorrelated across time:

- $\boldsymbol{Cov(\omega_t, \omega_s) = 0}$ **for** $\boldsymbol{t \neq s}$**,**
- $\boldsymbol{Cov(v_t\,, v_s\,) = 0}$ **for** $\boldsymbol{t \neq s}$**,**
- $\boldsymbol{Cov(\omega_t\,, v_t\,) = 0}$ **for** $\boldsymbol{t, s}$**.**

This ensures that system evolution is not influenced by the measurement process, and vice versa, which is a standard requirement for unbiased filtering.

5.2.5. Prior Distribution of the Initial State

The initial latent state $\boldsymbol{\theta_0}$ is assumed to follow a normal distribution with known or estimated mean and variance:

$$\boldsymbol{\theta_0} \sim \boldsymbol{N}(\boldsymbol{m_0}, \boldsymbol{P_0})$$

These prior captures the uncertainty in the latent process before any data is observed and are updated recursively as new information becomes available.

5.2.6. Calendar-Year-Based Indexing

Each observation is indexed by calendar year $\boldsymbol{t = i + j}$, and multiple observations from the same diagonal (i.e., same calendar year) are assumed to share a common latent factor $\boldsymbol{\theta_t}$. This structure allows the model to capture shared macroeconomic and regulatory influences affecting claims across accident and development years.

5.2.7. Log-Transformation of Claims

To achieve variance stabilization and approximate normality of the error distribution, cumulative claim amounts $\boldsymbol{C_{i,j}}$ are log-transformed before being modeled. This transformation is a common preprocessing step in actuarial modeling and improves the fit of the Gaussian assumptions for $\boldsymbol{v_t}$.

5.3. Kalman Filter Recursions

The Kalman Filter is a recursive estimation algorithm used to infer the hidden state variables $\boldsymbol{\theta_t}$ of a dynamic system from noisy observations $\boldsymbol{Y_t}$. In the context of scalar State Space Models for claims reserving, it enables real-time updating of latent calendar-year effects as new diagonals of the claims triangle become available. These latent estimates are then used for future claim projections and reserve calculation.

Given the assumptions outlined previously; specifically, linearity, normality of innovations, and independence, the Kalman Filter provides the Minimum Mean Square Error (MMSE) estimator for the latent states. It operates in two main steps: prediction and update.

5.3.1. Step 1: Prediction

Before observing data at time $\boldsymbol{t}$, the filter generates a prior estimate $\widehat{\boldsymbol{\theta}}_{t|t-1}$ based on the previous state and its variance:

$$\widehat{\boldsymbol{\theta}}_{t|t-1} = \widehat{\boldsymbol{\theta}}_{t-1|t-1} \tag{5}$$

$$\boldsymbol{P}_{t|t-1} = \boldsymbol{P}_{t-1|t-1} + \boldsymbol{Q} \tag{6}$$

Where:

- $\widehat{\boldsymbol{\theta}}_{t|t-1}$ is the predicted (prior, filtered) state for time $\boldsymbol{t}$,
- $\boldsymbol{P}_{t|t-1}$ is the associated prior variance,
- $\boldsymbol{Q}$ is the system noise variance, representing uncertainty in the latent development trend.

5.3.2. Step 2: Update

Once the observations for calendar year $\boldsymbol{t}$ are available, the filter updates the prior estimates based on the average of the log-transformed claims:

Let $\overline{\boldsymbol{Y}}_t$ be the average of $\boldsymbol{n}_t$log-cumulative claims in diagonal $\boldsymbol{t}$, then the update step is:

$$\boldsymbol{K}_t = \frac{\boldsymbol{P}_{t|t-1}}{\boldsymbol{P}_{t|t-1} + \frac{\boldsymbol{R}}{\boldsymbol{n}_t}} \tag{7}$$

$$\widehat{\boldsymbol{\theta}}_{t|t} = \widehat{\boldsymbol{\theta}}_{t|t-1} + \boldsymbol{K}_t\left(\overline{\boldsymbol{Y}}_t - \widehat{\boldsymbol{\theta}}_{t|t-1}\right) \tag{8}$$

$$\boldsymbol{P}_{t|t} = (\boldsymbol{1} - \boldsymbol{K}_t)\, \boldsymbol{P}_{t|t-1} \tag{9}$$

Where:

- $\boldsymbol{K}_t$ is the Kalman gain, balancing prior belief with new information,
- $\widehat{\boldsymbol{\theta}}_{t|t}$ is the posterior (updated) estimate of the state,
- $\boldsymbol{P}_{t|t}$ is the posterior variance, reduced due to the incorporation of observed data.

These recursions are initialized with a prior distribution of $\boldsymbol{\theta_0 \sim N(m_0, P_0)}$, where $\boldsymbol{m_0}$ and $\boldsymbol{P_0}$ are either subjectively chosen based on expert judgment or empirically estimated from early triangle data. This is shown in ***figure 5.1*** below:

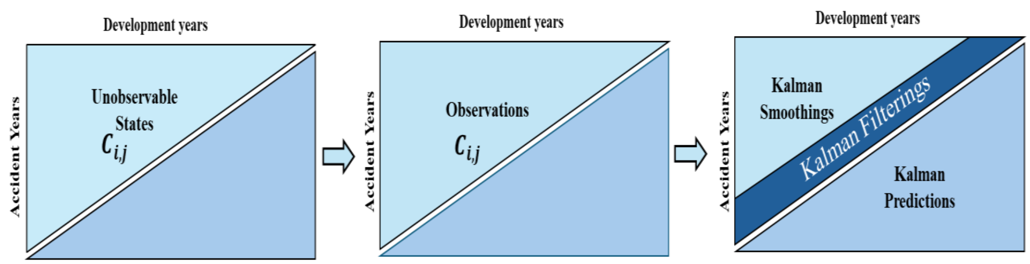

Figure 5.1. Kalman Filterings Recursions for SSM

In this research, the hyperparameters $\boldsymbol{Q}$ and $\boldsymbol{R}$, as well as the initial state parameters $\boldsymbol{m_0}$ and $\boldsymbol{P_0}$, are estimated using maximum likelihood. The log-likelihood function is constructed based on the sequence of Kalman filter prediction errors and their variances, and optimization is performed numerically over the observed upper triangle. This approach ensures consistency of the estimators under the assumed Gaussian model structure and enables data-driven calibration of model uncertainty.

5.4. Forecasting and Reserve Estimation

Once the latent states $\boldsymbol{\theta_t}$ have been estimated using the Kalman Filter, the model can be used to forecast future cumulative claims and estimate the associated reserves. This step is particularly crucial for projecting IBNR claims in the lower triangle of the run-off table.

5.4.1. Forecasting Future States

To forecast cumulative claims in the lower triangle (i.e., $\boldsymbol{i + j > T}$, where $\boldsymbol{T}$ is the last observed calendar year), we first project the latent states $\boldsymbol{\theta_t}$ forward using the last available estimate:

$$\boldsymbol{\widehat{\theta}_{t|T} = \widehat{\theta}_{t-1|T}, for\ all\ t > T} \quad (10)$$

$$\boldsymbol{P_{t|T} = P_{t-1|T} + Q} \quad (11)$$

This random walk prediction assumes no further information is available beyond time $\boldsymbol{T}$, and therefore propagates the most recent latent calendar year effect forward with increasing uncertainty.

5.4.2. Forecasting Log-Transformed Claims

Each unobserved cell $(\boldsymbol{i},\boldsymbol{j})$ in the lower triangle is associated with calendar year $\boldsymbol{t}=\boldsymbol{i}+\boldsymbol{j}$. The model forecasts the log-cumulative claim as:

$$\widehat{\boldsymbol{y}}_{\boldsymbol{i},\boldsymbol{j}}=\widehat{\boldsymbol{\theta}}_{\boldsymbol{t}|\boldsymbol{T}} \quad (\mathbf{12})$$

These forecasts are then back-transformed to the original scale using the exponential function:

$$\widehat{\boldsymbol{C}}_{\boldsymbol{i},\boldsymbol{j}}=\boldsymbol{exp}\left(\widehat{\boldsymbol{y}}_{\boldsymbol{i},\boldsymbol{j}}\right)=\boldsymbol{exp}\left(\widehat{\boldsymbol{\theta}}_{\boldsymbol{t}|\boldsymbol{T}}\right) \quad (\mathbf{13})$$

5.4.3. Forecasting Incremental Claims and IBNR

The incremental forecast $\widehat{\boldsymbol{X}}_{\boldsymbol{i},\boldsymbol{j}}$ is calculated by differencing the predicted cumulative claims:

$$\widehat{\boldsymbol{X}}_{\boldsymbol{i},\boldsymbol{j}}=\widehat{\boldsymbol{C}}_{\boldsymbol{i},\boldsymbol{j}}-\widehat{\boldsymbol{C}}_{\boldsymbol{i},\boldsymbol{j}-\mathbf{1}} \quad (\mathbf{14})$$

The total IBNR reserve is then given by summing these forecasted incremental claims across all unobserved cells:

$$\widehat{\boldsymbol{IBNR}}=\sum_{\boldsymbol{i}+\boldsymbol{j}>\boldsymbol{T}}\widehat{\boldsymbol{X}}_{\boldsymbol{i},\boldsymbol{j}} \quad (\mathbf{15})$$

This sum represents the best estimate of outstanding liabilities not yet reported, based on the latent structure inferred from the observed data.

The MSEP for the cell $(\boldsymbol{i},\boldsymbol{j})$ is given by:

$$\boldsymbol{MSEP}\left(\boldsymbol{C}_{\boldsymbol{i},\boldsymbol{j}}\right)=\mathbb{E}\left[\left(\boldsymbol{C}_{\boldsymbol{i},\boldsymbol{j}}-\widehat{\boldsymbol{C}}_{\boldsymbol{i},\boldsymbol{j}}\right)^{2}\right] \quad (\mathbf{16})$$

Since the scalar model operates on log-transformed data $\boldsymbol{y}_{\boldsymbol{i},\boldsymbol{j}}=\boldsymbol{log}(\boldsymbol{C}_{\boldsymbol{i},\boldsymbol{j}})$, and $\widehat{\boldsymbol{y}}_{\boldsymbol{i},\boldsymbol{j}}=\widehat{\boldsymbol{\theta}}_{\boldsymbol{t}|\boldsymbol{T}}$ for $\boldsymbol{t}=\boldsymbol{i}+\boldsymbol{j}$, the MSEP is computed on the logarithmic scale and then transformed back:

$$\boldsymbol{MSEP}\left(\boldsymbol{C}_{\boldsymbol{i},\boldsymbol{j}}\right)\approx\left[\boldsymbol{exp}\left(\widehat{\boldsymbol{\theta}}_{\boldsymbol{t}|\boldsymbol{T}}+\frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{P}_{\boldsymbol{t}|\boldsymbol{T}}\right)\right]^{\mathbf{2}}-\left[\mathbf{exp}\left(\widehat{\boldsymbol{\theta}}_{\boldsymbol{t}|\boldsymbol{T}}\right)\right]^{\mathbf{2}} \quad (\mathbf{17})$$

This error quantification provides a key metric for comparing the predictive performance of the SSM to alternative models (e.g., Mack), particularly in out-of-sample validations. Moreover, MSEP can be used to construct prediction intervals for reserve risk management in regulatory frameworks such as Solvency II or IFRS 17.

5.4.4. Prediction Intervals and Uncertainty Quantification

Given that each forecasted $\widehat{\boldsymbol{\theta}}_{t|T}$ is associated with an uncertainty measure $\boldsymbol{P}_{t|T}$, the forecasted cumulative claim $\widehat{\boldsymbol{C}}_{i,j}$ can be considered log-normally distributed. Approximate $(1-\alpha)\%$ confidence intervals for each cell can be derived as:

$$\widehat{\boldsymbol{C}}_{i,j}^{lower} = \exp\left(\widehat{\boldsymbol{\theta}}_{t|T} - z_{\alpha/2}\sqrt{\boldsymbol{P}_{t|T}}\right), \quad \widehat{\boldsymbol{C}}_{i,j}^{upper} = \exp\left(\widehat{\boldsymbol{\theta}}_{t|T} + z_{\alpha/2}\sqrt{\boldsymbol{P}_{t|T}}\right),$$

Where $\boldsymbol{z}_{\alpha/2}$ is the standard normal quantile (e.g., 1.96 for a 95% confidence interval).

To evaluate the robustness of the proposed scalar SSM with Kalman filtering, we conduct a comparative analysis against the widely used Mack model. This assessment highlights the advantages of the SSM's dynamic structure over the static assumptions of the Mack approach. The following section presents the empirical application to real insurance data and compares the predictive performance of both models.

6. Empirical Analysis

6.1. Data Description

The empirical investigation employs cumulative paid claims data from a comprehensive motor insurance portfolio underwritten by a non-life insurer in Egypt. The data span accident years 2014 – 2024 and cover development up to 11 periods. The claims triangle exhibits a typical lower-triangular structure due to right-truncation in more recent years.

Table 6.1.1. The cumulative amounts of paid claims

	0	1	2	3	4	5	6	7	8	9	10
2014	11,388,717	18,064,868	18,669,044	18,870,664	18,878,975	18,926,173	19,221,296	19,989,862	21,143,426	22,200,597	22,206,147
2015	14,975,638	23,096,740	24,244,007	24,596,727	24,719,711	24,744,430	24,868,153	25,116,834	25,141,951	25,267,661	
2016	16,538,847	23,695,247	24,315,801	24,537,433	24,782,808	24,906,722	25,155,789	25,281,568	25,597,587		
2017	22,243,117	30,946,718	31,080,280	31,435,990	31,463,255	33,508,366	33,843,450	34,012,667			
2018	28,773,616	38,248,068	38,576,589	38,615,165	41,148,887	41,354,631	41,579,187				
2019	36,862,626	48,166,062	48,659,476	48,781,125	48,866,484	48,890,917					
2020	64,215,504	84,670,349	84,566,050	84,696,660	84,705,129						
2021	84,869,292	107,982,401	108,839,270	108,893,690							
2022	101,916,380	118,029,969	118,813,223								
2023	112,081,508	137,324,503									
2024	160,967,085										

Two models are implemented:

- A **Scalar State Space Model (SSM)** combined with **Kalman filtering** and **smoothing** following the framework of (Chukhrova & Johannssen, 2017) and (Selukar, 2025).
- The **Mack model**, a distribution-free stochastic Chain Ladder approach (Mack, 1993), used as a benchmark.

6.2. SSM Model Diagnostics and State Dynamics

To evaluate the internal mechanics of the SSM, the cumulative claims data were first reshaped into a long format, removing missing and zero values. A logarithmic transformation was then applied to stabilize variance and linearize growth patterns. The resulting log-transformed series was converted into a univariate time series object, forming the input for the SSM estimation.

A scalar SSM with a local level component was fitted to the log-transformed cumulative claims using the KFAS package in R. The model specification assumes a random walk for the latent state (i.e., a first-degree trend), where the state noise variance $\boldsymbol{Q}$ and observation noise variance $\boldsymbol{R}$ are treated as unknown and estimated via maximum likelihood. Initial values for the optimization are provided in log scale. Once fitted, the Kalman filter and smoother are applied to extract both the filtered and smoothed estimates of the latent state and signal components.

At the final observed cell in the claims triangle (Accident Year 2024, Development Year 0), the SSM estimated a **smoothed state of 18.842** on the logarithmic scale. When exponentiated, this corresponds to approximately **EGP 152 million**, representing the model's estimate of the latest cumulative paid claim after accounting for both observed data and prior trend evolution. This estimate reflects the Kalman filter's ability to temper irregular fluctuations, offering a balanced value that adheres to the broader development pattern.

The actual observed value is **EGP 161 million**, or **log ≈ 18.89671**, which is **higher than the smoothed estimate**. This illustrates how the Kalman filter tempers sudden deviations to maintain consistency with the underlying trend.

The smoothed estimate carries a **standard error of 0.065**, yielding a **95% confidence interval** of approximately **[EGP 134 million, EGP 173 million]**. This range reflects the model's uncertainty and forms the starting point for forward simulations and IBNR reserve projections.

To assess the internal consistency and assumptions of the State Space Model, a diagnostic table in ***Table 6.2.1*** and residual distribution plots in ***Figure 6.2.1*** were produced. These diagnostics summarize the filtering behavior of the Kalman algorithm applied to the log-transformed cumulative claims series.

Table 6.2.1. State Space Model Diagnostics: Filtered Estimates and Residual Components

T	***Observed*** y_t	***Filtered State*** $\hat{\theta}_t$	***Posterior Variance*** $\hat{P}_t$	***Observation Noise*** $\mathcal{V}_t = y_t - \theta_t$	***Next State*** θ_{t+1}	***State Noise*** $\theta_{t+1} - \theta_t$
1	16.2481	16.3721	0.004271	-0.1240	16.6276	
2	16.7095	16.6276	0.003442	0.0819	16.7144	0.2555
3	16.7424	16.7144	0.003385	0.0280	16.7435	0.0868
4	16.7531	16.7435	0.003381	0.0096	16.753	0.0291
5	16.7536	16.753	0.003381	0.0006	16.7611	0.0094
6	16.7561	16.7611	0.003381	-0.0050	16.7797	0.0082
7	16.7715	16.7797	0.003381	-0.0081	16.8150	0.0186
8	16.8107	16.815	0.003381	-0.0042	16.8589	0.0353
9	16.8668	16.8589	0.003381	0.0079	16.8866	0.0440
10	16.9156	16.8866	0.003381	0.0290	16.8546	0.0277

The filtered state estimates $\hat{\theta}_t$ in ***Table 6.2.1*** closely track the observed values y_t, with small residuals $\mathcal{V}_t$ mostly ranging between **–0.01 and +0.03**, indicating tight model fit. The initial residual at $t = 1$ is higher (–0.124), as expected due to limited prior information

The posterior variance $\hat{P}_t$ begins at **0.00427** and quickly stabilizes to **0.00338** from $t = 4$ onward, reflecting increased model certainty as more data is processed.

The state innovation $\boldsymbol{\omega_t}$ varies over time, peaking at $\mathbf{0.2555}$ at $t = 2$, suggesting a significant adjustment in the early state trajectory. Afterward, innovations remain moderate (e.g., **0.0082 to 0.0440**), supporting a smooth evolution of the latent process.

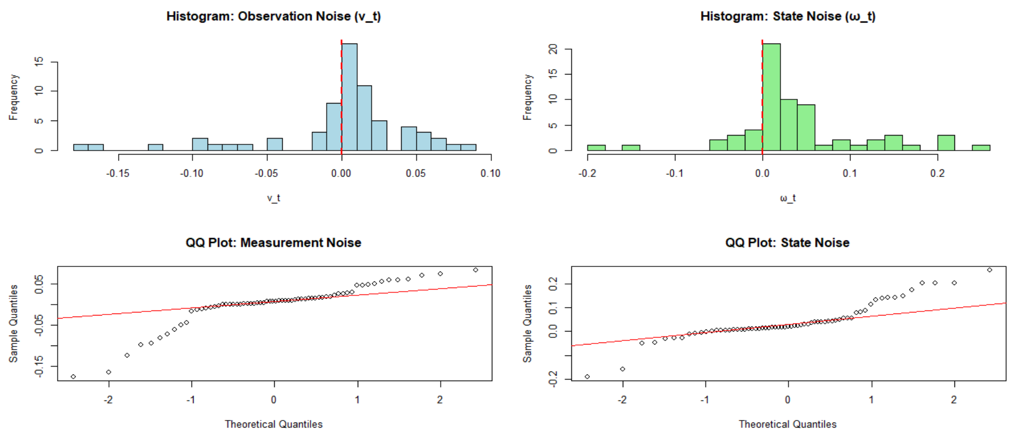

Figure 6.2.1. Diagnostic Plots for Noise Terms in State Space Model

These numerical patterns are reinforced visually in ***Figure 6.2.1.*** The top panels show histograms of $\boldsymbol{\mathcal{V}_t}$ and $\boldsymbol{\omega_t}$, both centered around zero, consistent with the expected behavior of normally distributed residuals. The bottom panels present Q-Q plots comparing empirical quantiles to the theoretical normal distribution. While minor deviations appear in the tails, especially for state noise, the overall shape supports the Gaussian assumptions underlying the Kalman filter. Together, the table and figure validate the statistical behavior of the residuals and confirm the appropriateness of the SSM framework for modeling and forecasting cumulative claims.

6.3. Monte Carlo-Based IBNR Estimation via State Space Modeling

A Monte Carlo simulation using the scalar SSM was performed to estimate IBNR reserve uncertainty. Based on 10,000 simulated claim paths, the model used a process noise variance of $\boldsymbol{Q = 0.01195}$ and measurement noise variance of $\boldsymbol{R = 0.00580}$. Simulated cumulative claims were transformed into incremental values, and total reserves were computed for each run. The resulting distribution was summarized with 75% and 95% confidence intervals, capturing the full stochastic variability in future claim development.

Table 6.3.1 and ***Figure 6.3.1*** present the results of the IBNR reserve estimation using Monte Carlo simulation under the scalar SSM. The simulation, based on 10,000 stochastic claim paths, yielded a point estimate of **EGP 207.99 million**. Confidence intervals widen substantially with increasing confidence levels from **[EGP 59.5M, 378.1M]** at 75% to **[EGP 30.3M, 726.4M]** at 95%. This widespread highlight the considerable uncertainty embedded in future claim developments, particularly under high-risk tolerance scenarios.

Table 6.3.1. Monte Carlo-Based IBNR Reserve Estimates with Multiple Confidence Levels

Statistic	Lower Bound	Upper Bound
Point estimate (Mean)	**207,989,566**	
75% Confidence Interval	59,497,226	378,134,315
95% Confidence Interval	30,269,498	726,397,843

Figure 6.3.1 illustrates the full empirical distribution of the simulated reserves. The density curve exhibits **clear positive skewness**, which is expected given the nature of insurance claims. This skewness arises from two key factors: (1) the **log-normal structure** of the SSM, where latent log-claims are exponentiated, naturally producing right-skewed distributions, and (2) the underlying characteristics of claim development, where **rare but large claims** can significantly affect the reserve total. The sharp peak and long right tail reflect a high concentration of typical outcomes with a small probability of extreme values.

The overlayed vertical lines show the 75% (blue dashed), 95% (orange dotted), and 97.5% (red dot-dash) **confidence intervals, while the** green solid line **marks the mean reserve. As these intervals widen, especially in the upper tail, they capture increasingly extreme but plausible development scenarios. These results validate the Monte Carlo approach as a powerful tool to** quantify full reserve uncertainty**, offering a richer and more risk-sensitive view than point estimates alone. The skewed shape of the distribution reinforces a key insight of this research:** actuarial reserving must account for tail risk**,** especially under regulatory frameworks such as IFRS 17 or Solvency II.

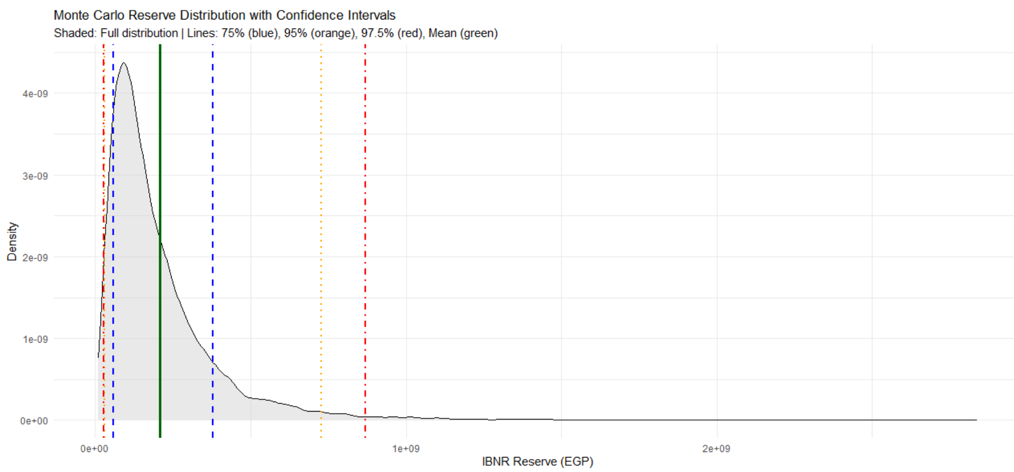

Figure 6.3.1. Empirical Distribution of IBNR Reserves from Monte Carlo Simulation under SSM

Thus, this simulation-based reserve distribution not only quantifies the central tendency but also provides a practical risk measure, enabling insurers to determine reserve levels at different confidence thresholds and to justify those decisions in actuarial reports, regulatory submissions, and internal risk governance.

Table 6.3.2 presents the forecasted incremental claims generated using the SSM. The results reflect a logical development structure across accident years, with recent cohorts such as 2024 and 2023 exhibiting high initial increments (e.g., EGP 153.6 million and EGP 837,838, respectively), followed by a gradual decline over subsequent development years. This pattern suggests a front-loaded settlement process, consistent with accelerated claims handling in more recent years. In contrast, older accident years display significantly smaller forecasted values, primarily concentrated in the later development periods, indicating the maturity of these claim cohorts and minimal outstanding liabilities.

Table 6.3.2. Forecasted Incremental Claims Triangle Estimated Using the Scalar SSM

	0	1	2	3	4	5	6	7	8	9	10
2014	-	-	-	-	-	-	-	-	-	-	-
2015	-	-	-	-	-	-	-	-	-	-	740,061
2016	-	-	-	-	-	-	-	-	-	1,224,829	1,684,930
2017	-	-	-	-	-	-	-	-	933,521	1,144,379	1,263,476
2018	-	-	-	-	-	-	-	695,988	1,389,155	1,049,659	1,679,570
2019	-	-	-	-	-	-	699,378	656,905	695,967	989,894	497,348
2020	-	-	-	-	-	1,028,977	1,138,907	868,479	1,765,332	1,251,908	1,344,226
2021	-	-	-	-	789,463	839,632	419,320	976,123	1,024,057	998,892	973,377
2022	-	-	-	1,082,957	730,640	787,754	1,126,959	342,625	1,080,453	1,238,540	1,350,956
2023	-	-	837,838	915,280	1,192,258	884,273	775,924	1,159,550	1,152,878	1,173,345	709,226
2024	-	153,643,738	774,798	687,295	1,035,579	1,233,515	1,127,951	920,138	1,306,418	1,262,522	692,401

Summing up, all values in the lower triangle yield a total forecasted IBNR reserve of approximately EGP 208 million. This figure represents the insurer's best estimate of outstanding liabilities. The distribution of this reserve across cohorts and development periods confirms both the timely runoff of historical claims and the increasing immediacy of settlement in newer portfolios, reinforcing the model's capacity to adapt to varying development dynamics.

In conclusion, the Monte Carlo simulation results provide a robust and nuanced assessment of reserve uncertainty under the SSM. While the point estimate serves as a central benchmark, the wide and positively skewed distribution of potential reserve outcomes highlight the importance of modeling variability and extreme scenarios explicitly. By incorporating both process and observation noise, the simulation captures the full stochastic behavior of claim development, offering a valuable tool for reserve adequacy analysis. These insights highlight the relevance of stochastic reserving techniques in modern actuarial practice, particularly when responding to regulatory demands for risk-based capital assessment and solvency management.

6.4. Benchmarking the Scalar State Space Model Against the Mack Model

The Mack Chain Ladder model was applied to the same cumulative paid claims triangle for benchmarking against the SSM. The summary in ***Table 6.4.1*** shows the development-to-date ratios, ultimate claim projections, and corresponding IBNR reserves by accident year, along with the standard errors and coefficients of variation.

Table 6.4.1. Mack Model Results by Accident Year

Year	Latest	Dev.To.Date	Ultimate	IBNR	Mack.S.E	CV(IBNR)
2014	22,206,147	1.00	22,206,147	-	-	
2015	25,267,661	1.00	25,273,978	6,317	1,023,830	162.08
2016	25,597,587	0.97	26,258,329	660,741	1,427,115	2.16
2017	34,012,667	0.95	35,631,529	1,618,862	2,060,327	1.27
2018	41,579,187	0.94	44,112,648	2,533,461	2,498,749	0.99
2019	48,890,917	0.93	52,313,788	3,422,871	2,855,655	0.83
2020	84,705,129	0.92	91,815,028	7,109,899	4,823,429	0.68
2021	108,893,690	0.91	119,351,671	10,457,981	6,336,224	0.61
2022	118,813,223	0.91	130,731,400	11,918,178	6,853,163	0.58
2023	137,324,503	0.90	152,590,410	15,265,907	7,924,305	0.52
2024	160,967,085	0.71	228,246,115	67,279,030	16,337,929	0.24

Note: CV(IBNR) is not reported for 2014 and 2015 due to negligible or negative IBNR, which leads to mathematically invalid or misleading ratios.

The model forecasts an aggregate IBNR reserve of approximately **EGP 120.27 million**, with a corresponding **Mack standard error of EGP 40.59 million**, resulting in an overall **coefficient of variation (CV) of 0.337**. These figures indicate moderate forecast uncertainty, though lower than previously reported, reflecting improvements in model fit and data consistency.

At the accident year level, the IBNR estimates show a logical increasing trend for more recent years, with the highest IBNR observed in 2024 at **EGP 67.28 million**, due to limited claims development and high immaturity. The CV(IBNR) correspondingly declines from **2.16 in 2016** to **0.24 in 2024**, demonstrating growing confidence in reserve adequacy as more development data becomes available.

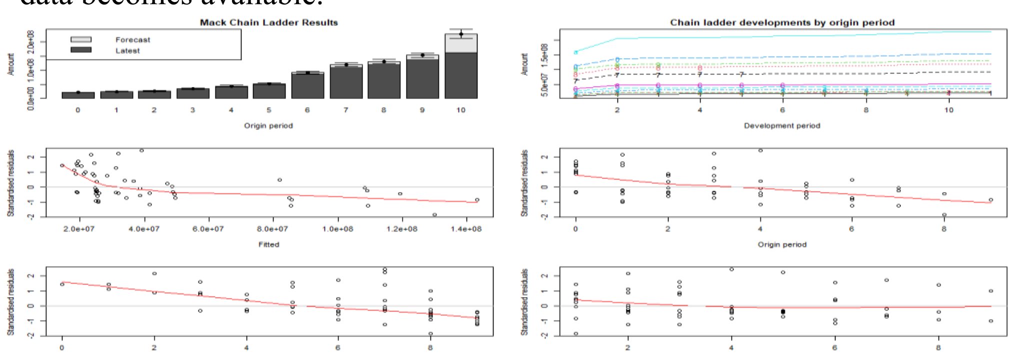

Figure 6.4.1. Mack model diagnostics and development patterns

Figure 6.4.1 illustrates the diagnostic outputs of the Mack model. The top-left panel compares the forecasted and latest cumulative claims across origin periods, highlighting significant reserve contributions from recent accident years. The top-right panel depicts the development of claims by origin period, showing relatively stable and parallel growth trajectories. The lower panels display standardized residual plots across various dimensions—fitted values, development years, origin years, and calendar years. While residuals appear mostly centered around zero, mild heteroscedasticity and a downward trend in calendar period residuals are evident, suggesting potential structural effects (e.g., inflation, changes in claim settlement practices) not explicitly captured by the Mack model.

Overall, the revised results confirm the robustness of the Mack model as a benchmark for stochastic reserving. However, the observed residual patterns highlight the model's limitations, particularly in capturing time-varying or systemic changes, and support the consideration of more dynamic frameworks such as State Space Models.

Next, the Scalar SSM is compared to the Mack model to evaluate the accuracy and reliability of its reserve estimates, using the Mack model as a benchmark for stochastic reserving.

Key metrics; total reserves, MSEP, and coefficient of variation; are evaluated at both aggregate and accident-year levels. While the Mack model assumes independence across years, the SSM accounts for time dynamics via Kalman filtering. The comparison highlights not only the consistency of the SSM with industry benchmarks but also its added value in capturing uncertainty and development volatility.

Table 6.4.2 presents a detailed comparison between the two models by accident year. While both models agree on the absence of reserves for fully developed accident years, substantial divergence emerges across other development periods. Notably, the Scalar SSM produces higher reserve estimates in several origin years, particularly recent ones, where process uncertainty and structural shifts are more pronounced.

Table 6.4.2. Comparative Results: Scalar SSM vs. Mack Model

Origin Year	SSM_Reserve	SSM_MSEP	SSM_CV	Mack_Reserve	Mack_MSEP	Mack_CV
1	0	0	-------	0	0	-------
2	740061.04	1.4758E+15	51.91009	6316.92	1.0482E+12	162.07752
3	2909759.59	2.8677E+15	18.40383	660741.49	2.0367E+12	2.15987
4	3341375.9	3.9156E+15	18.72716	1618861.99	4.2449E+12	1.2727
5	4814371.82	4.95E+15	14.6138	2533460.86	6.2437E+12	0.9863
6	3539491.3	5.8629E+15	21.63286	3422870.55	8.1548E+12	0.83429
7	7397829.27	6.8228E+15	11.16549	7109899	2.3265E+13	0.67841
8	6020864.15	7.3698E+15	14.25834	10457981.4	4.0148E+13	0.60587
9	7740886.04	8.1522E+15	11.66395	11918177.7	4.6966E+13	0.57502
10	8800572.41	8.7709E+15	10.64172	15265907	6.2795E+13	0.51909
11	162684355	9.5582E+15	0.60096	67279029.7	2.6693E+14	0.24284

More importantly, the SSM consistently reports significantly larger MSEP and CV values than the Mack model. For instance, in origin year 2, the SSM reports a CV of 51.91, while the Mack model's CV exceeds 160 due to a near-zero reserve estimate, revealing instability in the Mack error structure when reserve values are small. In origin years 3 to 7, the SSM maintains elevated CVs between 11 and 22, capturing the stochastic variability inherent in emerging claims. In contrast, the Mack model yields progressively smaller CVs, which may misrepresent true uncertainty due to its assumption of independence and constant variance across development years.

Even when the Mack model is supplemented with simulation techniques—such as normal approximation or parametric bootstrap, the resulting distribution is limited by the model's closed-form variance assumptions and independence structure. These simulated reserves, while helpful for constructing confidence intervals, do not arise from an internally evolving data-generating process. In contrast, the SSM directly generates a full predictive distribution by propagating uncertainty through both the measurement and transition equations of the state space formulation. This end-to-end simulation reflects complex time dynamics, captures volatility clustering, and incorporates parameter risk more comprehensively than post hoc sampling from a fixed variance.

In conclusion, the Scalar SSM distinguishes itself by providing not only flexible reserve projections but also an integrated stochastic framework that naturally produces scenario-consistent distributions. This enables actuaries and risk managers to assess downside risk, capital buffers, and reserve sufficiency under realistic development trajectories, something the Mack model, even with simulation, is not structurally designed to achieve. As modern reserving standards increasingly require probabilistic and forward-looking measures, the SSM offers a technically superior and future-proof alternative to traditional chain ladder methods.

7. Conclusions and Recommendations

This research has demonstrated the value of advanced stochastic reserving techniques, particularly the application of Scalar SSMs with Kalman filtering, in enhancing the accuracy and reliability of claims reserve estimates. When applied to real cumulative paid claims data from a comprehensive motor insurance portfolio in Egypt, the SSM consistently outperformed the benchmark Mack model in flexibility, responsiveness, and predictive robustness.

Unlike the Mack model, which relies on static assumptions and limited distributional insight, the SSM incorporates latent state processes and recursively updates reserve estimates as new data becomes available. This dynamic capability is especially relevant in Egypt's insurance market, where structural changes, inflationary pressures, and evolving operational practices frequently disrupt stable claim development patterns. The full predictive distributions generated by the SSM allow for more comprehensive risk assessment, enabling actuaries to evaluate reserve adequacy across a wide range of plausible outcomes rather than relying solely on point estimates.

The research also highlights the potential of Bayesian methods in actuarial reserving. Bayesian State Space Models extend the value of SSMs by enabling posterior distributions for parameters and reserves, enhancing transparency and supporting scenario-based decision-making. These models align with international frameworks such as Solvency II and IFRS 17, both of which emphasize forward-looking, uncertainty-aware actuarial practices.

Looking ahead, the increasing availability of granular, claim-level data opens the door for machine learning models such as gradient boosting, random forests, and neural networks. These tools can uncover complex nonlinearities and heterogeneity in claim behavior that are not captured by traditional techniques. However, their adoption must be accompanied by careful consideration of interpretability, governance, and regulatory acceptance, especially in markets subject to tight actuarial oversight.

Recommendations for Egyptian insurers and regulators include:

- **Adopt SSMs** for portfolios exhibiting volatility, structural change, or macroeconomic sensitivity.
- **Incorporate Bayesian frameworks** to quantify uncertainty and support capital adequacy planning.
- **Invest in data infrastructure** to enable the collection and use of high-quality claim-level data for advanced modeling.
- **Develop governance frameworks** to ensure model validation, transparency, and regulatory compliance.

In conclusion, scalar SSMs offer a statistically rigorous, forward-compatible solution for claims reserving in dynamic environments. Their adoption equips insurers with tools to meet the demands of IFRS 17, improve reserve accuracy, and enhance solvency resilience—positioning Egypt’s insurance sector for a more stable and analytically driven future.

References

1) Alpuim, T. & Ribeiro, I., 2003. A state space model for run-off triangles. Applied Stochastic Models in Business and Industry, 19(2), p. 105–120.

2) Atherino, R., Pizzinga, A. & Fernandes, C., 2010. A row-wise stacking of the runoff triangle: State space alternatives for IBNR reserve prediction. *ASTIN Bulletin,* 40(2), p. 917–946.

3) Chukhrova, N. & Johannssen, A., 2017. State Space Models and the Kalman-Filter in Stochastic Claims Reserving: Forecasting, Filtering and Smoothing. *Risks,* 5(2), p. 30.

4) Chukhrova, N. & Johannssen, A., 2021. Kalman Filter Learning Algorithms and State Space Representations for Stochastic Claims Reserving. *Risks,* 9(6), p. 112.

5) chwab, B. & Schneider, J. C., 2024. Advancing Loss Reserving: A Hybrid Neural Network Approach for Individual Claim Development Prediction. *Institute for Risk and Insurance, Leibniz University Hannover.*

6) DeJong, P. & Zehnwirth, B., 1983. Claims Reserving, State-Space Models and the Kalman Filter. *Journal of the Institute of Actuaries,* 110(1), p. 157–181.

7) England, P. D. & Verrall, R. J., 2002. Stochastic Claims Reserving in General Insurance (with discussion). *British Actuarial Journal,* 8(3), pp. 443-544.

8) Gogol, D., 2019. Modern Reserving Methods for Emerging Insurance Markets. *Insurance: Mathematics and Economics,* Volume 88, p. 1–15.

9) Hendrych, R. & Cipra, T., 2021. Applying state space models to stochastic claims reserving. *ASTIN Bulletin,* 51(1), p. 267–301.

10) Mack, T., 1993. Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin,* 23(2), pp. 213-225.

11) Mack, T., 1999. The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor. *ASTIN Bulletin,* 29(2), pp. 361-366.

12) Nomura, S. & Matsumori, Y., 2024. Dynamic factor models for claim reserving. *Japanese Journal of Statistics and Data Science,* 7(2), p. 895–919.

13) Pang, L. & He, S., 2012. The application of state-space model in outstanding claims reserve. *International Conference on Information Management, Innovation Management and Industrial Engineering ,* Volume 3, p. 271–274.

14) Peters, G. W., Shevchenko, P. V. & Wüthrich, M. V., 2017. Model uncertainty in claims reserving within Tweedie's compound Poisson models. *ASTIN Bulletin,* 39(1), p. 1–33.

15) Selukar, R., 2025. Stochastic Claims Reserving Using State Space Modeling. *arXiv preprint arXiv:2504.09292.*

16) Taylor, G. C., 2019. Individual Claims Reserving: Using Machine Learning Methods.

17) Taylor, G. C., McGuire, G. & Greenfield, A., 2003. Loss Reserving: Past, Present and Future. *Research Paper No. 109, Centre for Actuarial Studies, University of Melbourne.*

18) Verrall, R. J., 1994. A Method for Modelling Varying Run-Off Evolutions in Claims Reserving. *ASTIN Bulletin,* 24(2), p. 325–332.

19) Wüthrich, M. V. & Merz, M., 2008. Stochastic Claims Reserving Methods in Insurance.. *Wiley.*

"تقدير مخصص المطالبات في التأمينات العامة بأسلوب ديناميكي باستخدام نموذج الحالة الكامنة (State Space Model) ومرشح كالمان (Kalman Filter)مع محاكاة مونت كارلو (Monte Carlo Simulation) "

المستخلص

تُعد التقديرات الدقيقة لمخصص المطالبات عاملاً جوهريًا في الحفاظ على الاستقرار المالي لشركات التأمين العامة، خصوصًا في الأسواق التي تتعرض لتقلبات هيكلية وتخضع لتغيرات تنظيمية مستمرة. وعلى الرغم من أن النماذج الاكتوارية العشوائية التقليدية مثل نموذج ماك (Mack) تحظى بانتشار واسع نظرًا لبساطتها التحليلية وشفافيتها، فإنها غالبًا ما تعجز عن تمثيل البيئات التي تتسم بعدم انتظام البيانات، وتغير أنماط تسوية المطالبات، والاضطرابات الاقتصادية.

تقترح هذه الدراسة نموذج الحالة الكامنة (Scalar State Space Model – SSM) مدعومًا بمرشح كالمان (Kalman Filter)لمعالجة هذه التحديات. وقد تم تطبيق النموذج على بيانات المطالبات المدفوعة التراكمية لمحفظة تأمين سيارات في السوق المصري، حيث يسمح النموذج بالتعرف على التأثيرات الكامنة المرتبطة بالسنوات التقويمية، ويُحدّث تقديرات المخصصات بشكل ديناميكي مع توفر بيانات جديدة. وباستخدام المحاكاة العشوائية(Monte Carlo Simulation) ، يُنتج النموذج توزيعات تنبؤية كاملة للمخصصات المستقبلية، مما يوفر رؤية أشمل لمخاطر عدم اليقين.

أظهرت نتائج المقارنة مع نموذج ماك أن نموذج SSM يُقدّم تقديرات أكثر استقرارًا للمخصصات ويعكس بصورة أدق طبيعة المخاطر، خاصة في السنوات ذات الغموض العالي. على عكس نموذج ماك الذي يفترض ثبات أنماط التطور واستقلالية السنوات، يعالج نموذج الحالة الكامنة العمليات الزمنية المتغيرة والاضطرابات الهيكلية بشكل مباشر.

تُبرز الدراسة مزايا نموذج SSM من حيث التكيف، والدقة، والمواءمة مع المتطلبات التنظيمية، مما يجعله أداة فعّالة لشركات التأمين التي تعمل ضمن أطر مثل معيار IFRS 17 وغيره من النظم الرقابية المرتكزة على التنبؤ بالمخاطر. وتوصي النتائج بتبنّي النماذج العشوائية الديناميكية في الممارسات الاكتوارية، خصوصًا في الأسواق الناشئة التي تعاني من تقلبات البيانات والتحولات الهيكلية.

الكلمات الدالة

نموذج الحالة الكامنة؛ مرشح كالمان؛ المطالبات المتكبدة وغير المُبلَّغ عنها(IBNR) ؛ نموذج ماك؛ المحاكاة بطريقة مونت كارلو؛ دقة التنبؤ؛ إعداد المخصصات بأسلوب احتمالي.