

# Modellvalidierung – VaR-Modell

André Miemiec<sup>1</sup>, Kerstin Steinberg

FRAME Consulting GmbH

Gabriel-Max-Straße 12

10245 Berlin,

## **Abstract**

In diesem Artikel soll das allgemeine Vorgehen einer Modellvalidierung, wie es in dem Artikel ‚Modellvalidierung – Eine Blaupause‘ definiert wurde [1], am Beispiel der Validierung eines „Value at Risk“-Modells (VaR) konkretisiert werden. Das Herzstück der Ausdifferenzierung des allgemeinen Vorgehens einer Modellvalidierung bildet dabei die Formulierung eines für das Validierungsobjekt spezifischen Vorgehensmodells. Ein solches wird hier für die Validierung eines VaR-Modells detailliert beschrieben und verprobt. Eine gesonderte Diskussion wird sich mit der Anwendbarkeit des Wurzelgesetzes zur Ermittlung des VaR für einen vorgegebenen Risikohorizont beschäftigen.

---

<sup>1</sup> Corresponding author: [andre.miemiec@frame-consult.de](mailto:andre.miemiec@frame-consult.de)

# 1 Einleitung

Die Validierung der Berechnung des Value at Risk (VaR) für ein Portfolio ist ein einfacher aber keineswegs trivialer Vorgang, da er mannigfaltigste Informationen letztlich zu einem einzigen Skalar verdichten muß. Während der Verdichtung werden exakte Ergebnisse zum Teil mit Faustformeln vermengt, um das finale Ergebnis zu produzieren<sup>2</sup>. Dies macht eine vollumfängliche Validierung zu einem aufwendigen Vorgang, weil insbesondere die Angemessenheit dieser Faustformeln kritisch zu hinterfragen ist.

Eine sinnvolle Validierung eines VaR-Modells kann aber auch mit einem überschaubaren Aufwand erreicht werden. Der Schlüssel dafür ist die Auswahl eines geeigneten Vorgehensmodells, das die wichtigsten, idealerweise möglichst disjunkten Einzelschritte der VaR-Erzeugung gründlich durchleuchtet. Durch die richtige Auswahl und anschließende Analyse derjenigen Einzelschritte, die das Wesen der Analytik des VaR-Modells repräsentieren, werden die typischsten Fehlerquellen eines VaR-Modells ausgeschaltet und die Wahrscheinlichkeit quasi auf null reduziert, daß sich schwerwiegende fachliche Fehler im VaR-Modell eingeschlichen haben.

Ein wichtiger Aspekt ist dabei auch die organisatorisch saubere Trennung von Risiko- und Bewertungsmethodik.

Die Atome der Risikomethodik sind die Risikofaktoren, die der Bewertungsmethodik die marktkonformen Bewertungsmodelle. Neben den vorhandenen Überschneidungen dieser Atome existieren aber auch deutliche Unterschiede. Es ist gerade ihre systematische Differenzierung, die dazu führt, ein „Faß ohne Boden“ bei den erforderlichen Analysen zu vermeiden.

Im Rahmen der Risikomethodik wird also primär die Frage beantwortet, ob alle Risikofaktoren korrekt erfaßt sind. Die Frage nach der fachlich korrekten Abbildung der Bewertungsfunktion ist dann aber wiederum Aufgabe der Bewertungsvalidierung. Die Risikovalidierung wird somit auch zu einem Trigger für die Bewertungsvalidierung<sup>3</sup>.

Der Punkt der Validierung von historischen Zeitreihen für die Zwecke der Risikorechnung wird in diesem Artikel bewußt ausgeklammert, weil die Darstellung der erforderlichen Techniken den Rahmen der Illustration der Konkretisierung des allgemeinen Validierungsvorgehens [1] auf den Anwendungsfall der Validierung eines VaR-Modells sprengen würde. Dieser essentielle Qualitätssicherungsprozeß wird daher im Folgenden als erledigt angesehen.

---

<sup>2</sup> Faustformeln kommen z.B. bei der Skalierung des VaR auf einen vorgegebenen Risikohorizont bzw. der Aggregation unterschiedlicher Risikoarten zum Einsatz.

<sup>3</sup> Die Frage der ggf. nötigen Vereinfachung eines marktkonformen Soll-Modells für die Zwecke einer adäquaten Risikoberechnung (im Sinne eines Gleichgewichts zwischen fachlichem Gehalt und Performanceanforderungen) ist dann wiederum eine Aufgabe der Risikomethodik. Wie die Überschneidung der beiden Themen praktisch gehandhabt werden können, wird in den später folgenden Anmerkungen zum Backtesting des VaR-Modells deutlich.

Für die Lesbarkeit des Artikels wird die Vertrautheit mit den Grundkonzepten von Value-at-Risk-Modelle vorausgesetzt.

Der Rest des Artikels ist folgendermaßen gegliedert. In Abschnitt 2 wird zunächst das gewählte Vorgehensmodell zur Validierung des VaR vorgestellt. In Abschnitt 3 wird dieses Vorgehensmodell verprobt, indem die Validierung einer VaR-Berechnung am Beispiel des Tools JSONrisk der FRAME Consulting GmbH durchgeführt wird [8]. Praktisch wird die Validierung damit durch einen Abgleich der VaR-Rechnung im Zielsystems (hier JSONrisk) mit den Ergebnissen der VaR-Berechnung eines Referenzsystems (hier Excel) durchgeführt. Abschnitt 4 faßt die wesentlichen Aussagen dieses Artikels zusammen.

## 2 Modellvalidierungsprozeß VaR

Dieser Abschnitt stellt das zur Validierung eines VaR-Modells gewählte Vorgehensmodell vor und zeigt, weshalb die dadurch erreichte Modelldurchleuchtung ausreichend ist, um die Qualität eines VaR-Modells zu beurteilen.

### 2.1 Vorgehensmodell

Das Vorgehensmodell zur Validierung des VaR-Modells wird aus drei Teilschritten bestehen, die in Abbildung 1 gezeigt sind und im Folgenden näher beschrieben werden. Pro Teilschritt muß dabei ein Akzeptanzkriterium festgelegt werden, das es erlaubt zu beurteilen, ob ein Validierungsschritt erfolgreich abgeschlossen wurde.

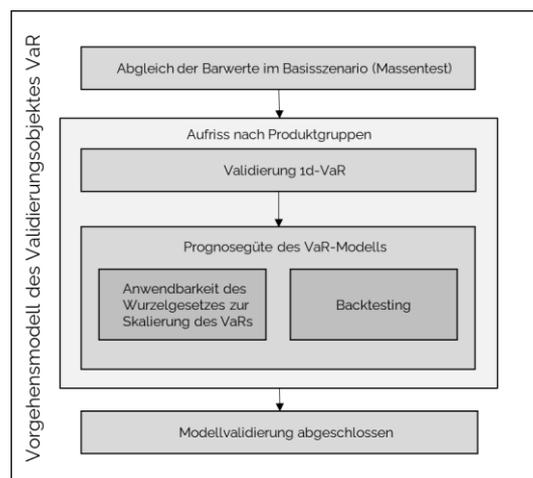


Abbildung 1: Vorgehensmodell zur Validierung des VaR-Modells.

#### Teilschritte:

1. Der erste Schritt besteht aus dem Abgleich der Barwerte des Gesamtportfolios zwischen dem Ziel- und dem zur Validierung benutzten Referenzsystem im BASE-Szenario. Dieser dient primär der Sicherstellung der Vergleichbarkeit der Barwerte zwischen Ziel- und Referenzsystem.

Akzeptanzkriterium: Das Akzeptanzkriterium dieses Schrittes ist eine maximale prozentuale Abweichung der beiden Portfoliobarwerte. Der Wert dieser Abweichung wird z.B. auf 1 % festgelegt. Die genaue Höhe dieser Schranke ist dabei zweitrangig, solange sie die sinnvolle Vergleichbarkeit der beiden Portfolioabbildungen gewährleistet.

2. Der zweite Schritt beinhaltet die Validierung des 1d-VaR auf ausgewählten (Teil-)Portfolien (vgl. hierzu auch die erläuternden Kommentare am Ende dieser Aufzählung).

Akzeptanzkriterium: Als Akzeptanzkriterium für den Abgleich der (Teil-)Portfolien wird eine maximale relative Abweichung der 1d-VaRs zwischen dem Ziel- und dem Referenzsystem von 5% angesetzt. Die Höhe dieser Abweichung orientiert sich an der Schranke, die für die Modellvalidierung von internen Modellen vorgesehen ist (vgl. [6]).

3. Im dritten Schritt wird die Prognosegüte des VaR-Modells auf (Teil-)Portfolien untersucht. Dazu werden mindestens die beiden folgenden Aspekte untersucht:

1. Prüfung der Normalverteilungsannahme, d.h. Anwendbarkeit der Wurzelskalierung zur Bestimmung des Value at Risk für einen gegebenen Risikohorizont<sup>4</sup>. Bei der Wurzelskalierung handelt es sich um eine typische Faustformel, die ein verteilungsannahmefreies VaR-Modell mit einem auf einer spezifischen Verteilungsannahme basierenden VaR-Modell vermengt.

In der Praxis ist deshalb selbst bei Verwendung einer historischen Simulation eine Prüfung der Normalverteilungsannahme erforderlich, wenn das Wurzelgesetz angewendet werden soll. Diese erfolgt qualitativ mittels QQ-Plots und quantitativ mittels eines statistischen Tests (z.B. Kolmogorov-Smirnov-Tests).

Akzeptanzkriterium: Das Akzeptanzkriterium für die qualitative Prüfung auf Normalverteilung ist, daß der QQ-Plot<sup>5</sup> optisch keine groben Abweichungen von einer Geraden aufweist.

Das Akzeptanzkriterium für die quantitative Prüfung auf Normalverteilung ist die Bestätigung der Null-Hypothese des statistischen Tests (vgl. hierzu Abschnitt 5.1 im Anhang).

Das Gesamturteil ergibt sich aus der logischen Und-Verknüpfung des quantitativen und des qualitativen Teilurteils.

2. Backtesting<sup>6</sup>: Bestimmung der Prognosegüte des VaR-Modells durch Ausreißerzählung.

---

<sup>4</sup> Für die Anwendbarkeit der Wurzelskalierung müssen die P&L-Veränderungen auf dem täglichen und dem Zielhorizont näherungsweise normalverteilt sein. Ist dies nicht der Fall, erfolgt die Übertragung des 1d-VaR auf einen nd-VaR durch eine n-fache Faltung, wobei n die Anzahl der Tage des Risikohorizonts bezeichnet.

<sup>5</sup> Ein QQ-Plot stellt die Quantile einer empirischen Verteilung den Quantilen (z.B.) einer Normalverteilung gegenüber. Gleichheit der Verteilungen impliziert eine 45°-Gerade im QQ-Plot.

<sup>6</sup> Eine Unterscheidung zwischen einem Clean- und einem Dirty-Backtesting wird hier nicht durchgeführt, da sie für die grundsätzlichen Überlegungen in Kombination mit dem in Abschnitt 3.1 festgelegten Beispielportfolio nicht erforderlich ist. In der Praxis ist diese Differenzierung erforderlich.

Akzeptanzkriterium: Das Akzeptanzkriterium ist, daß der VaR die regulatorische Ampel einhält (vgl. hierzu Abschnitt 5.2, Tabelle 2). Das Kriterium wird sowohl auf den VaR als auch auf den negativen VaR angewendet, weil auch ein Auftauchen zu häufiger positiver P&L-Überziehungen auf einen Defekt des Modells hinweisen kann.

### Kommentare:

- i. Der zweite und der dritte Teilschritt des Vorgehensmodells verwendet eine Iteration über ausgewählte Produktgruppen (Teilportfolios). Im Fall eines einfachen Interest-Rate-Portfolios genügt es z.B., die in Abbildung 2 gezeigten Produktgruppen zu untersuchen. Die nötigen Produktgruppen werden in einem realistischen Fall durch das Produktspektrum des konkreten Bankportfolios bestimmt und so sortiert, daß eine möglichst vollständige Nachbildung des VaRs auch auf Einzelgeschäften durchgeführt werden kann.

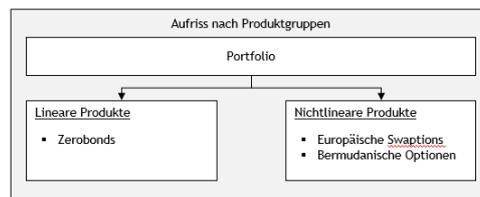


Abbildung 2: Iteration der Teilschritte 2 und 3 des Vorgehensmodells über Produktgruppen.

- ii. Anschließend müssen der zweite und dritte Teilschritt noch einmal auf dem Gesamtportfolio durchgeführt werden, um die Korrektheit der Aggregation zu prüfen. Bei der Analyse von VaR-Modellen, die auf historischer Simulation beruhen, ist dieser letzte Schritt aber eher ein Funktionstest als eine Validierung, weil die Aggregation nur aus einer Zusammenführung von P&L-Vektoren besteht und keine wesentliche fachliche Komponente mehr besitzt. Bei klassischen Varianz-Kovarianz-Ansätzen für den VaR ist die Aggregation jedoch eine zentrale Validierungsthematik, weil erst hier die Korrelationen modellseitig ihre volle Wirkung entfalten.
- iii. Im Fall der Zusammenführung von VaRs für unterschiedliche Risikoarten (z.B. Marktpreisrisiko und OpRisk) ist eine Analyse der Einzelfalllösung für die Aggregation erforderlich. Diese Thematik wird hier aus Platzgründen nicht betrachtet und ggf. in einem zukünftigen Beitrag thematisiert.

## 2.2 Dokumentationstemplate Teilschritt 3

Zur einheitlichen Darstellung der Ergebnisse bei der Durchführung der produktgruppenweisen Validierung wird eine standardisierte Darstellung der Ergebnisse gewählt. Diese besteht aus zwei Teilen.

1. Test der Normalverteilungsannahme:
  - a. Qualitativ: Graphische Darstellung der täglichen P&L-Verteilung als Histogramm sowie der zugehörige QQ-Plot der empirischen Verteilung (vgl. Abbildung 3).

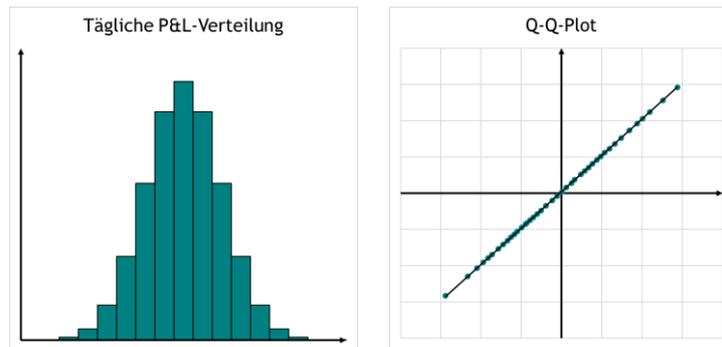


Abbildung 3: Prototypische Darstellung von P&L-Verteilung und QQ-Plot

- b. Quantitativ: Ergebnis des statistischen Tests auf Normalverteilung.
2. Backtesting: Zählung der positiven (grün) und negativen (rot) Ausreißer der täglichen P&L-Veränderungen gegenüber dem Niveau des positiven bzw. negativen Betrags des VaRs (blau).  
Die negativen Ausreißer stellen eine Modellverletzung im regulatorischen Sinne dar. Die zusätzliche Betrachtung der positiven Ausreißer ist ein ergänzendes Mittel zur Prüfung der Modellkonsistenz.

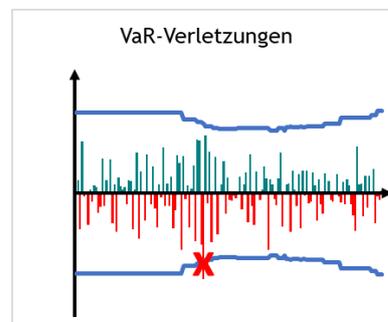


Abbildung 4: Prototypische Darstellung der Ausreißerzählung.

### 3 Validierung eines VaR-Modells

Das im Abschnitt 2 vorgestellte Vorgehensmodell zur Validierung eines VaR-Modells soll in diesem Abschnitt an einem konkreten Beispiel veranschaulicht werden. Das Validierungsobjekt ist der Value at Risk, der mittels einer historischen Simulation basierend auf den letzten 252 Handelstagen erzeugt wird.

#### 3.1 Hypothetisches Bankportfolio

Zu Illustrationszwecken wird ab hier ein vereinfachtes Portfolio betrachtet, das aus genau 3 Einzelprodukten besteht, die stellvertretend für die in Abbildung 2 genannten Teilportfolios stehen.

Konkret handelt es sich dabei per Stichtag um

- einen Zerobond mit Maturity in 2 Jahren,
- eine Europäische Receiver Swaption (1x5) und um
- eine Bermudanische Bondoption (5x10) mit einer Call-Frequenz von 6M.

Es wird zusätzlich angenommen, daß in der hypothetischen Bank die Optionen lediglich über die Verwendung von ATM-Volatilitäten abgebildet worden sind.

**Kommentar:** Die hier gemachten Annahmen stellen keine wesentlichen Restriktionen für die Anwendbarkeit des Vorgehensmodells dar<sup>7</sup>. Insbesondere läßt sich das Modell um das hier abgegrenzte Thema der Validierung der Risikozeitreihen unproblematisch erweitern.

## 3.2 Durchführung der Validierung

### 3.2.1 Teilschritt 1

Der erste Schritt des Vorgehensmodells sieht den Abgleich des Barwertes im BASE-Szenario zwischen Referenzsystem und Zielsystem vor. In unserem Beispiel stimmten diese im Ziel- und Referenzsystemen hinreichend überein. Eine Abweichung entsteht lediglich aus der Bermudanischen Bondoption und diese bewegt sich unterhalb der gesetzten Schranke von 1% und ist auf marginale Implementierungsunterschiede der beiden verwendeten Hull-White-Modelle zwischen Ziel- und Referenzsystem zurückzuführen.

### 3.2.2 Teilschritt 2

Die Replikation des 1d-VaRs des Zielsystems im Referenzsystem liefert für das betrachte Portfolio vergleichbare Ergebnisse, die innerhalb des für diesen Teilschritt formulierten Akzeptanzkriteriums liegen. Auf den exakt nachrechenbaren Fällen des Zerobonds und der Europäischen Swaption lagen die Unterschiede sogar innerhalb der numerischen Toleranzen.

### 3.2.3 Teilschritt 3

#### 3.2.3.1 *Bankportfolio*

Für das Risiko des Bankportfolios sind die beiden Risikoarten ‚Zins‘ und ‚Volatilität‘ maßgeblich.

Die folgenden graphischen Darstellungen der täglichen P&L-Verteilung und des entsprechenden QQ-Plots legen nahe, daß die empirische Verteilung der täglichen P&Ls des Bankportfolios einer Normalverteilung entspricht. Kleinere Abweichungen von der Normalverteilung existieren nur an den Rändern der Verteilung.

---

<sup>7</sup> Für ein realistisches Bankportfolio kann man die Zerlegung in Teilportfolien über beliebige Tiefen bis auf Einzelgeschäfte durchführen. In der Regel genügen drei Tiefen: Bankportfolio, Teilportfolio, Einzelgeschäft.

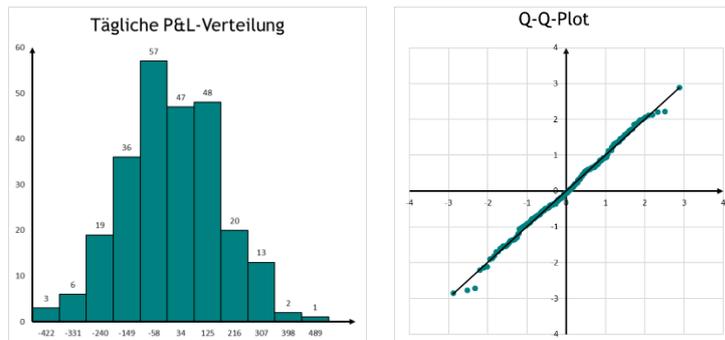


Abbildung 5: Verteilung der tägliche P&Ls sowie der zugehörige QQ-Plot.

Die Teststatistik für den Hypothesentest auf Normalverteilung liefert einen Wert, der kleiner als der charakteristische Wert des Kolmogorov-Smirnov-Tests ist (s. Abschnitt 5.1). Somit wird die Nullhypothese, daß die empirische Verteilung einer Normalverteilung entspricht, nicht verworfen.

Abbildung 6 zeigt die Zahl der Ausreißer im Backtesting.

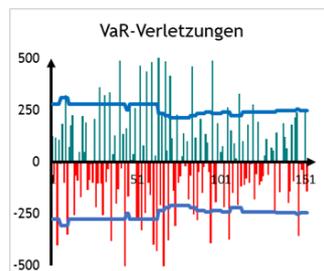


Abbildung 6: Einschätzung der Prognosegüte des VaR-Modells durch Ausreißerzählung.

Der Abbildung ist zu entnehmen, daß die negativen P&L-Änderungen (rot) 17-mal die durch den VaR gesetzte untere Schranke und daß die positiven P&L-Änderungen (grün) 17-mal die durch den VaR gesetzte obere Schranke reißen. Das entspricht einer roten Ampel (vgl. Abschnitt 5.2, Tabelle 2).

### Analyse:

Die Analyse für die Quellen dieser hohen Zahl an Ausreißern wird im Folgenden durch einen Aufriß nach den relevanten Produktgruppen durchgeführt. Dort wird gezeigt, daß die schlechte Prognosegüte des VaR-Modells zum Großteil auf die unvollständige Berücksichtigung einer Risikoart (,Volatilität') zurückzuführen ist.

### Wertung:

Die QQ-Plots und die Teststatistik stützen die Normalverteilungsannahme. Das Backtesting führt zu einer roten Ampel (vgl. Abschnitt 5.2, Tabelle 2). Die Analyse für die Ursache des Auftauchens einer roten Ampel erfolgt in den folgenden produktgruppen-spezifischen Abschnitten.

#### 3.2.3.2 Zerobond

Bei der Bewertung eines Zerobonds spielt nur die Risikoart ,Zins' eine Rolle.

Die folgenden graphischen Darstellungen der täglichen P&L-Verteilung und des entsprechenden QQ-Plots legen nahe, daß die empirische Verteilung der täglichen

P&Ls eines Zerobonds einer Normalverteilung entspricht. Kleinere Abweichungen von der Normalverteilung existieren nur an den Rändern der Verteilung.

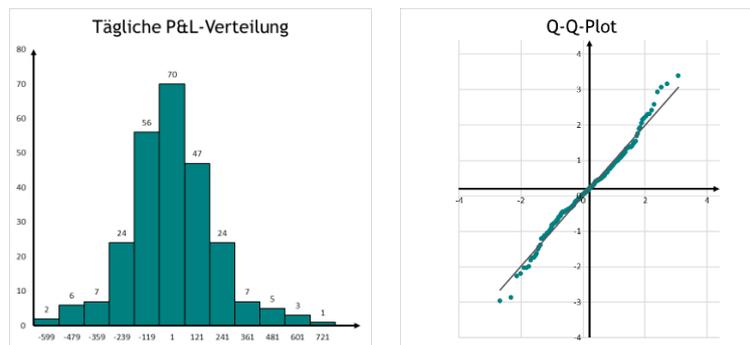


Abbildung 7: Verteilung der tägliche P&Ls sowie der zugehörige QQ-Plot.

Die Teststatistik für den Hypothesentest auf Normalverteilung liefert einen Wert, der kleiner als der charakteristische Wert des Kolmogorov-Smirnov-Tests ist (s. Abschnitt 5.1). Somit wird die Nullhypothese, daß die empirische Verteilung einer Normalverteilung entspricht, nicht verworfen.

Abbildung 8 zeigt die Zahl der Ausreißer im Backtesting.

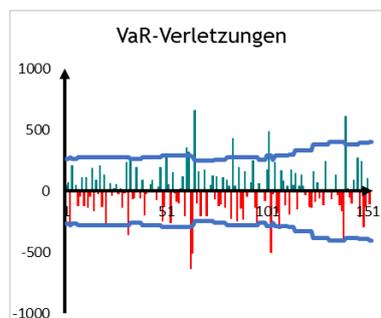


Abbildung 8: Einschätzung der Prognosegüte des VaR-Modells durch Ausreißerzählung.

Der Abbildung ist zu entnehmen, daß die negativen P&L-Änderungen (rot) 8-mal die durch den VaR gesetzte untere Schranke und daß die positiven P&L-Änderungen (grün) 7-mal die durch den VaR gesetzte obere Schranke reißen. Das entspricht einer grünen Ampel (vgl. Abschnitt 5.2, Tabelle 2).

### Analyse:

Nicht erforderlich.

### Wertung:

Die QQ-Plots und die Teststatistik stützen die Normalverteilungsannahme. Das Backtesting führt zu einer grünen Ampel (vgl. Abschnitt 5.2, Tabelle 2).

### 3.2.3.3 Europäische Swaption

Bei der Bewertung einer Europäischen Swaption, tritt die Risikoart ‚Volatilität‘ neu hinzu.

Die folgenden graphischen Darstellungen der täglichen P&L-Verteilung und des entsprechenden QQ-Plots legen nahe, daß die empirische Verteilung der täglichen P&Ls einer Europäischen Swaption einer Normalverteilung entspricht. Kleinere Abweichungen von der Normalverteilung existieren nur an den Rändern der Verteilung.

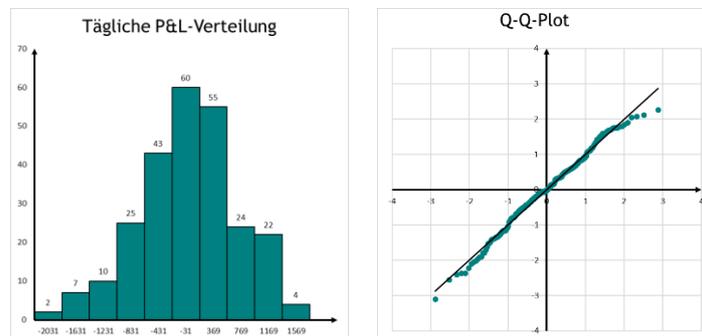


Abbildung 9: Verteilung der tägliche P&Ls sowie der zugehörige QQ-Plot.

Die Teststatistik für den Hypothesentest auf Normalverteilung liefert einen Wert, der kleiner als der charakteristische Wert des Kolmogorov-Smirnov-Tests ist (s. Abschnitt 5.1). Somit wird die Nullhypothese, daß die empirische Verteilung einer Normalverteilung entspricht, nicht verworfen.

Die folgende Abbildung 10 zeigt die Zahl der Ausreißer im Backtesting<sup>8</sup>.

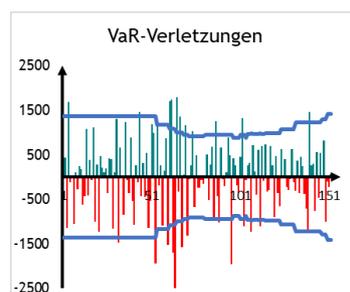


Abbildung 10: Einschätzung der Prognosegüte des VaR-Modells durch Ausreißerzählung.

Der Abbildung ist zu entnehmen, daß die negativen P&L-Änderungen (rot) 13-mal die durch den VaR gesetzte untere Schranke und daß die positiven P&L-Änderungen (grün) 12-mal die durch den VaR gesetzte obere Schranke reißen. Im Sinne der Ampel (vgl. Abschnitt 5.2, Tabelle 2) weist das auf ein mögliches Problem des Risikomodells hin (gelbe Ampel).

### Analyse:

Die weitere Analyse zeigt, daß ein Grund für das schlechtere Abschneiden des Teilportfolios die eingangs gemachte Annahme der Verwendung von ATM-Volatilitäten ist. Die Risikoart ‚Volatilität‘ besitzt neben der Unterausprägung ‚ATM-

<sup>8</sup> Die linken Flanken der beiden durch den VaR gesetzten Begrenzungen für die P&L sind konstant (blaue Kurven in Abbildung 10). Dies wurde manuell geprüft und als ein Artefakt einer tatsächlich vorliegenden Ausreißerkonstellation für wahr erkannt.

Risiko' (hier abgebildet) auch die Unterausprägung ‚Smile-Risiko‘ (hier nicht abgebildet). Bestimmt man den VaR mit einem Modell für europäische Optionen, das den Smile berücksichtigt, dann ergibt sich das Bild aus Abbildung 11.

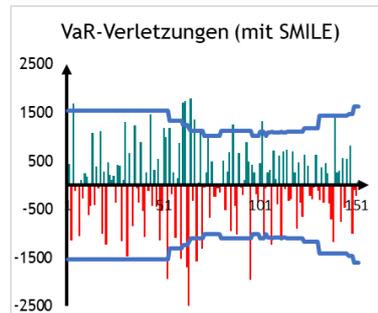


Abbildung 11: Einschätzung der Prognosegüte des VaR-Modells durch Ausreißerzählung (SMILE-basiert).

Bei Berücksichtigung des Smiles ist der Abbildung zu entnehmen, daß die negativen P&L-Änderungen (rot) 10-mal die durch den VaR gesetzte untere Schranke und daß die positiven P&L-Änderungen (grün) 10-mal die durch den VaR gesetzte obere Schranke reißen. Das entspräche einer grünen Ampel (vgl. Abschnitt 5.2, Tabelle 2).

### Wertung:

Die QQ-Plots und die Teststatistik stützen die Normalverteilungsannahme. Das Backtesting führt zu einer gelben Ampel (vgl. Abschnitt 5.2, Tabelle 2).

Die Analyse zeigte, daß die schlechtere Prognosegüte des VaR-Modells auf diesem Teilportfolio auf die unvollständige Berücksichtigung einer Risikoart (‚Volatilität‘) zurückzuführen ist. Eine Berücksichtigung des Smiles im VaR-Modell würde zu einer grünen Ampel im Backtesting führen (vgl. Abschnitt 5.2, Tabelle 2).

#### 3.2.3.4 Bermudanische Bondoption

Bei der Bewertung Bermudanischer Bondoptionen, spielt ebenso wie bei Europäischen Swaptions die Volatilität eine zentrale Rolle.

Die folgenden graphischen Darstellungen der täglichen P&L-Verteilung und des entsprechenden QQ-Plots legen nahe, daß die empirische Verteilung der täglichen P&Ls einer Bermudanischen Bondoption einer Normalverteilung entspricht. Kleinere Abweichungen von der Normalverteilung existieren nur an den Rändern der Verteilung.

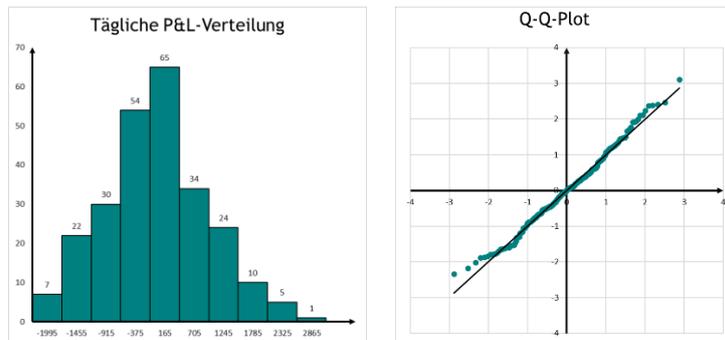


Abbildung 12: Verteilung der tägliche P&Ls sowie der zugehörige QQ-Plot.

Die Teststatistik für den Hypothesentest auf Normalverteilung liefert einen Wert, der kleiner als der charakteristische Wert des Kolmogorov-Smirnov-Tests ist (s. Abschnitt 5.1). Somit wird die Nullhypothese, daß die empirische Verteilung einer Normalverteilung entspricht, nicht verworfen.

Abbildung 13 zeigt die Zahl der Ausreißer im Backtesting.

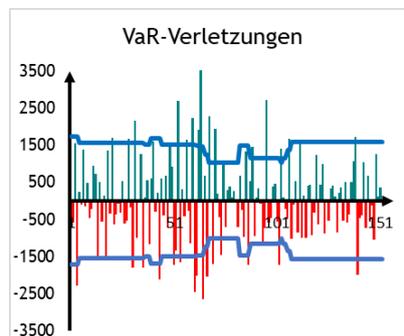


Abbildung 13: Einschätzung der Prognosegüte des VaR-Modells durch Ausreißerzählung.

Der Abbildung ist zu entnehmen, daß die negativen P&L-Änderungen (rot) 16-mal die durch den VaR gesetzte untere Schranke und daß die positiven P&L-Änderungen (grün) 16-mal die durch den VaR gesetzte obere Schranke reißen.

Im Sinne der Ampel (vgl. Abschnitt 5.2, Tabelle 2) weist das auf ein wahrscheinliches Problem des Risikomodells hin (gelbe Ampel an der Grenze zur roten Zone).

### Analyse:

Die weitere Analyse zeigt, daß das schlechte Abschneiden des Teilportfolios den gleichen Grund hat, der bereits bei der Europäischen Swaption ermittelt wurde. Bestimmt man den VaR mit einem Modell für Bermudanische Bondoptionen, das den Smile berücksichtigt, dann ergibt sich das Bild aus Abbildung 14.

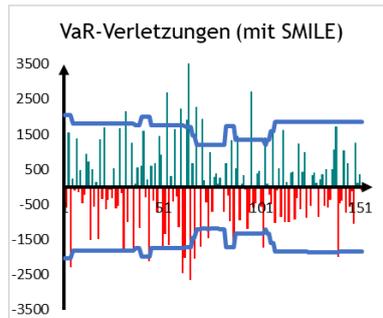


Abbildung 14: Einschätzung der Prognosegüte des VaR-Modells durch Ausreißerzählung (SMILE-basiert).

Bei Berücksichtigung des Smiles ist der Abbildung zu entnehmen, daß die negativen P&L-Änderungen (rot) 12-mal die durch den VaR gesetzte untere Schranke und daß die positiven P&L-Änderungen (grün) 11-mal die durch den VaR gesetzte obere Schranke reißen.

### Wertung:

Die QQ-Plots und die Teststatistik stützen die Normalverteilungsannahme. Das Backtesting führt zu einer gelben Ampel, die sich knapp vor der roten Zone befindet (vgl. Abschnitt 5.2, Tabelle 2).

Die Analyse zeigte, daß die schlechte Prognosegüte des VaR-Modells auf diesem Teilportfolio auf die unvollständige Berücksichtigung einer Risikoart (‚Volatilität‘) zurückzuführen ist. Eine Korrektur des Modells würde zu einer gelben Ampel im Backtesting führen, die sich knapp vor der grünen Zone befindet (vgl. Abschnitt 5.2, Tabelle 2).

## 3.3 Ergebnisse der Validierung

Die hier durchgeführte Validierung zeigte, daß das zur Illustration konstruierte VaR-Modell eine deutliche Schwäche beim Backtesting der volatilitätsabhängigen Produkte aufwies.

**Findings:** Diese Auffälligkeit konnte mit der unvollständigen Abbildung des Volatilitätssmile in Verbindung gebracht werden (vgl. Annahmen zur Volatilität aus Abschnitt 3.1).

**Maßnahme:** Die Maßnahme ist in diesem Fall die Auslösung einer Bewertungsmodellvalidierung mit dem Auftrag ein marktübliches Modell zu bestimmen, das die Grundlage einer zukünftigen VaR-Berechnung für die betroffenen Produktgruppen bildet.

**Auswirkung auf den Modellvalidierungsprozeß:** Die Auswirkung auf den Modellvalidierungsprozeß ist eine Revalidierung des VaR-Modells nach Abschluß der entsprechenden Modelländerungen.

## 4 Zusammenfassung

An dem hier präsentierten Beispiel der Validierung eines VaR-Modells konnten bereits einige typische Validierungshandlungen illustriert werden.

Neben der Feststellung zur Unvollständigkeit der Abbildung der Risikoart ‚Volatilität‘, die z.B. im Rahmen des Backtestings transparent gemacht werden konnte, ist bei der Validierung der Frage nachgegangen worden, inwieweit die Mischung einer historischen Simulation mit den Faustformeln, die auf einer Normalverteilungsannahme beruhen, gerechtfertigt werden kann.

Im Einzelnen:

- Das Backtesting der VaR-Modelle ist ein nützliches Werkzeug, um Marktinkonformitäten der Risikoabbildung aufzudecken. Die tägliche P&L berücksichtigt per se alle relevanten Risikofaktoren, das VaR-Modell aber nur die explizit modellierten Risikofaktoren. Das Verfehlen einer ‚grünen‘-Ampel beruht häufig auf einer unvollständigen Abbildung der Risiken, so daß das VaR-Modell einen zu geringen VaR prognostiziert<sup>9</sup>. Hierbei sollte aus Analysesicht aber auf jeden Fall die prozessuale Trennung zwischen der Risikomodellierung und der Bewertungsvalidierung beachtet werden, um ein ‚Faß ohne Boden‘ zu vermeiden.
- Die Validierung der Normalverteilungsannahme durch einen statistischen Hypothesentest ist naheliegend. Dann ist aber die Wirkweise des angewendeten Tests hinreichend zu untersuchen. Der VaR ist ein Quantilsmaß, das durch den Tail bestimmt wird. Die Analysen der QQ-Plots aus Abschnitt 3.2 machten deutlich, daß die fachlich interessanten Fragen mit den Tails der Verteilung verbunden sind („fat tails“). Aus der statistischen Literatur ist das Problem der (ungenügenden) Tail-Sensitivität eines statistischen Tests bekannt. Auch der Kolmogorov-Smirnov-Test besitzt diese Eigenschaft<sup>10</sup>, so daß das Bestehen des statistischen Tests noch keine endgültige Aussage über die Anwendbarkeit des Wurzelgesetzes zur Übertragung der täglichen P&L-Verteilung auf einen beliebigen Risikohorizont erlaubt. In der Folge kann es daher trotz eines erfolgreichen Hypothesentests durch die Anwendung der Wurzelskalierung zu einer falschen Risikoeinschätzung kommen.

---

<sup>9</sup> Bei der Zählung der Ausreißer ist immer noch die Anzahl, der durch Marktevents ausgelösten Ausreißer separat zu betrachten. Diese Gruppe ist mit den üblichen Grundvoraussetzungen eines VaR-Modells per se unverträglich.

<sup>10</sup> Es ist bekannt, daß der Kolmogorov-Smirnov-Test keine gute Sensitivität bzgl. Abweichung von der Normalverteilung in den Tails hat. Entsprechend wurden Ansätze für ‚tail‘-sensitivere Kolmogorov-Smirnov-Tests z.B. in [2] diskutiert.

Um dieses Problem auszuschließen, wurde die Übertragung der täglichen P&L-Verteilungen auf einen beliebigen Risikohorizont durch Faltungen der täglichen Verteilungen bestimmt. Das Ergebnis dieser Analyse ist in Abbildung 15 gezeigt. Sie bestätigt, daß für das betrachtete Beispielportfolio die Anwendbarkeit des Wurzelgesetzes angemessen ist.

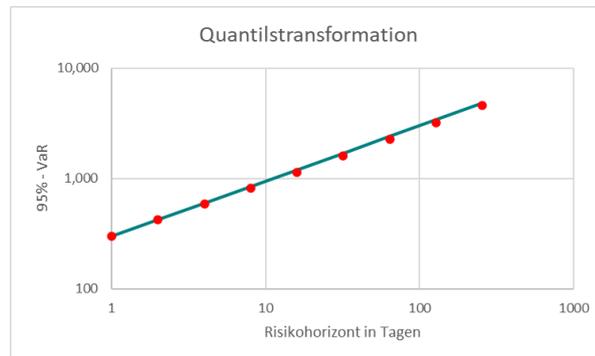


Abbildung 15: Quantilstransformationen via Faltung (rote Punkte) und via Wurzelskalierung (grüne Linie)

Eine mögliche Abweichung vom Wurzelgesetz würde sich in einem anderen Anstieg einer durch die roten Punkte gelegten Regressionsgeraden ausdrücken. Dann ließe sich das Skalierungsgesetz aber in die Form

$$VaR_{\alpha}(n\text{Tagen}) = n^{1/2+\beta} \cdot VaR_{\alpha}(1\text{Tag})$$

bringen lassen. Die Kalibrierung des zusätzlichen Parameters  $\beta$  trüge dann der Abweichung vom Wurzelgesetz Rechnung<sup>11</sup>.

In diesem Artikel wurden insbesondere zwei wichtige Themen aus Platzgründen ausgespart:

1. Die Thematik der Zeitreihenvalidierung wurde ausgeklammert, weil die Darstellung der erforderlichen Techniken den Zweck der Anwendbarkeit des allgemeinen Validierungsvorgehens nach [1] auf den Anwendungsfall der Validierung eines VaR-Modells unnötig verkompliziert hätte.
2. Die Thematik der Aggregation von VaR-Zahlen, die zu unterschiedlichen Risiken gehören (z.B. Marktpreisrisiko und OpRisk), wurde ebenfalls ausgeklammert, weil sie sich aufgrund der Einfachheit des hier betrachteten Bankportfolios nicht sinnvoll hätte diskutieren lassen.

Beide Themen werden ggf. in gesonderten Artikeln behandelt werden.

Ein abschließender Kommentar in eigener Sache betrifft die Anwendbarkeit der frei verfügbaren und quellenoffenen JSONrisk Library für die Zwecke einer VaR-Validierung. Auch wenn im Rahmen dieses Artikels nur ein hypothetisches Portfolio aus drei Geschäften betrachtet wurde, stellt die Portfoliogröße für die JSONrisk-Library keine Einschränkung dar. In dem Artikel [3] wurde ein Studium des

<sup>11</sup> Ein offener Punkt in dieser Analyse bleibt die Validierung der Unabhängigkeitsannahme der Zufallsvariablen, die durch die individuellen täglichen P&L-Verteilungen beschrieben werden (vgl. hierzu z.B. [7]). Diese wird sowohl für die Skalierung als auch für die Faltung vorausgesetzt.

Skalierungsverhaltens der Laufzeit von JSONRisk für ein Portfolio von Bermudanischen Optionen – eines der weitaus laufzeitintensivsten Produkte in realistischen Bankportfolien – durchgeführt und gezeigt, daß für handelsübliche Laptops gute und im Falle der Verwendung von AWS Lambda (Cluster) sogar konkurrenzlose Laufzeiten bei der VaR-Berechnung erreicht werden. Durch die zusätzliche Architektur der Library, die die sensiblen Daten ausschließlich auf dem Banksystem vorhält, ist sie ein ideales Validierungswerkzeug für die Durchführung performanceintensiver Vergleichsrechnungen.

## Literatur

- [1] A. Miemiec, "Modellvalidierung – Eine Blaupause", (2021), Available at SSRN: <https://ssrn.com/abstract=3987155>
- [2] D. M. Mason, J. H. Schuenemeyer, "A modified Kolmogorov-Smirnov-Test sensitive to tail alternatives", *The Annals of Statistics*, (1983), Vol. 11, No. 3
- [3] K. Steinberg, „Bewertung von großen Optionsportfolien mit JSONRisk“, (2021), [https://frame-consult.de/docs/Bewertung\\_von\\_grossen\\_Optionsportfolien.pdf](https://frame-consult.de/docs/Bewertung_von_grossen_Optionsportfolien.pdf)
- [4] S. D. Campbell, "A Review of Backtesting and Backtesting Procedures", Finance and Economics Discussion Series, Federal Reserve, (2005)
- [5] BCBS, "Supervisory Framework for the use of 'Backtesting' in Conjunction with the Internal Models Approach to Market Risk Capital Requirements", (1996)
- [6] EBA/RTS/2014/10, "RTS on Model Extensions and Changes for Market Risk", (2014)
- [7] P. Christoffersen, "Evaluating interval forecasts". *International Economic Review*, (1998), 39 (4), 841-862.
- [8] FRAME Consulting GmbH. JSONRisk-Bibliothek, <https://www.jsonrisk.de>

## 5 Anhang

### 5.1 Kolmogorov-Smirnov-Test

In Abschnitt 3 wird der Kolmogorov-Smirnov-Test verwendet, um die Nullhypothese ( $H_0$ ) zu prüfen, daß die empirische Verteilung der täglichen P&Ls mit einer Normalverteilungsannahme verträglich ist. Für ein vorgegebenes Signifikanzniveau  $\alpha$  (Konfidenzniveau  $1 - \alpha$ ) erfolgt die Prüfung der Nullhypothese üblicherweise durch den Vergleich der Kolmogorov-Smirnov-Teststatistik mit einem kritischen Wert. Der kritische Wert des Kolmogorov-Smirnov-Tests zu einem gegebenen Signifikanzniveau für eine Anzahl von  $n$ -Beobachtungen kann statistischen Tafeln entnommen werden<sup>12</sup>.

Der kritische Wert zum Signifikanzniveau 5% und einer Stichprobengröße von  $n=252$  Beobachtungen entspricht z.B. näherungsweise  $\sqrt{-0.5 \cdot \ln(\alpha/2)} / \sqrt{n} = 0.086$ . Die

---

<sup>12</sup> L. H. Miller, "Table of Percentage Points of Kolmogorov Statistics", *J. Am. Stat. Assoc.* 51:273, 111-121, 1956.

Ergebnisse des Kolmogorov-Smirnov-Tests für die 4 Teilportfolien aus Abschnitt 3 sind in Tabelle 1 zusammengefaßt.

Portfolio	Teststatistik	H <sub>0</sub> behalten
Bankportfolio	0.036	Ja
1. Teilportfolio (Zerobond)	0.060	Ja
2. Teilportfolio (Europäische Swaption)	0.040	Ja
3. Teilportfolio (Bermudanische Bondoption)	0.041	Ja

Tabelle 1: Ergebnisse der Teststatistik des Kolmogorov-Smirnov-Tests für die betrachteten Teilportfolien

## 5.2 Regulatorische Ausreißerampel

Die Zahl von Ausreißern im Backtesting ist durch den folgenden stochastischen Prozeß definiert [4]:

$$Z(\alpha, n) = \sum_{i=1}^n \mathbb{I}_{P\&L(1d) \geq VaR_{\alpha}(1d)}$$

$n$  bezeichnet hier die Länge der Backtesting-Periode in Tagen. Ist das VaR-Modell in sich schlüssig, dann treten einzelne Überziehungen des 1d-VaR mit der Wahrscheinlichkeit  $\alpha$  auf. Insgesamt erhält man somit für die Zahl der Ausreißer eine Binomialverteilung. Die Wahrscheinlichkeit für die Beobachtung von mehr als  $k$  Ausreißern läßt sich exakt berechnen und bildet die Grundlage des Baseler Ampelsystems [5]). Dieses baut auf der Wahrscheinlichkeit für das Begehen eines Fehlers erster Art (Ablehnung eines korrekten Modells) auf.

Die Grenzen für die Ampelsystematik sind über die folgenden beiden Wahrscheinlichkeiten für das Begehen eines Fehlers erster Art festgelegt:

- Rot: Wahrscheinlichkeit für mehr als  $k$  Ausreißer ist kleiner als 0.01%
- Gelb: Wahrscheinlichkeit für mehr als  $k$  Ausreißer ist größer als 0.01% aber kleiner als 5%
- Grün: Sonst.

Für die hier verwendeten Parameter einer VaR-Zeitreihenlänge von 152 Tagen und einem Konfidenzniveau von 95% ergeben sich die folgenden Zonengrenzen:

Ampelfarbe	#Ausreißer
Grün	< 12
Gelb	≥12 und < 17
Rot	≥ 17

Tabelle 2: Regulatorische Ampelsystematik des Backtestings

Der Vollständigkeit halber sei hier noch angemerkt: Verwendet man eine VaR-Zeitreihenlänge von 250 Tagen und ein Konfidenzniveau von 99%, dann bekommt man über die gleiche Methodik die in der Praxis geläufigen Grenzen für die Zonen von 5 bzw. 10 Ausreißern.

[F]