

Heep-Altiner, Maria; Münchow, Philipp; Scuzzarello, Vanessa

## Research Report

# Ausgleichsrechnungen mit Gauß Markow Modellen am Beispiel eines fiktiven Stornobestandes

Forschung am iwvKöln, No. 8/2014

## Provided in Cooperation with:

Technische Hochschule Köln – University of Applied Sciences, Institute for Insurance Studies

*Suggested Citation:* Heep-Altiner, Maria; Münchow, Philipp; Scuzzarello, Vanessa (2014) :  
Ausgleichsrechnungen mit Gauß Markow Modellen am Beispiel eines fiktiven Stornobestandes,  
Forschung am iwvKöln, No. 8/2014, Technische Hochschule Köln, Institut für Versicherungswesen  
(iwvKöln), Köln,  
<https://nbn-resolving.de/urn:nbn:de:hbz:832-cos-682>

This Version is available at:

<https://hdl.handle.net/10419/226566>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc-nd/3.0/>

**Forschung am IVW Köln, 8/2014**

Institut für Versicherungswesen

# **Ausgleichsrechnungen mit Gauß Markow Modellen am Beispiel eines fiktiven Stornobestandes**

**Maria Heep-Altiner, Philipp Münchow,  
Vanessa Scuzzarello**



**Maria Heep-Altiner, Philipp Münchow, Vanessa Scuzzarello**  
**Forschungsstelle FaRis**

## **Ausgleichsrechnungen mit Gauß Markow Modellen am Beispiel eines fiktiven Stornobestandes**

---

### **Zusammenfassung**

Die lineare Regression ist das bekannteste Verfahren zur Fehlerausgleichung, welches relativ einfach umgesetzt werden kann. Verallgemeinerte lineare Modelle sind eine zweckmäßige Erweiterung dieses Verfahrens, können aber aufgrund ihrer hohen Komplexität i. d. R. nur mit spezieller Software gerechnet werden. Aus diesem Grund wird in dieser Arbeit anhand eines fiktiven Storno-bestandes illustriert, wie man Ausgleichsrechnungen auch mit Hilfe von Gauß Markow Modellen durchführen kann, die mittels EXCEL und Visual Basic ebenfalls noch vergleichsweise einfach umsetzbar sind.

### **Abstract**

Linear Regression is a well-known algorithm for error calibration being relatively easy to execute. Generalized linear models define a suitable enlargement of this algorithm. Normally, they can only be performed with special software packages according to their high complexity. Therefore, it will be illustrated in this paper how to carry out an error equalization for virtual lapse data by Gauss Markov models being relatively easy to implement in EXCEL and Visual Basic.

### **Schlagwörter:**

Ausgleichsverfahren, Gauß Markow Modelle, Stornoanalyse, Versicherungsmathematik, Versicherungswissenschaft

# Inhaltsverzeichnis

1	VORBEMERKUNGEN .....	1
1.1	ALLGEMEINER ANSATZ EINER LINEAREN REGRESSION.....	1
1.2	GEWICHTETE MODELLE.....	2
1.3	MULTIPLIKATIVE MODELLE .....	2
1.4	CREDIBILITY MODELLE .....	3
2	GAUß MARKOW MODELLE .....	4
2.1	ALLGEMEINER ANSATZ EINES GAUß MARKOW MODELLS.....	4
2.2	GEWICHTETE MODELLE.....	5
2.3	QUALITATIVE MERKMALE.....	6
2.3.1	<i>Additiver Modellansatz</i> .....	7
2.3.2	<i>Multiplikativer Modellansatz</i> .....	7
2.4	TECHNISCHE UMSETZUNG MIT HILFE DER CHOLESKY ZERLEGUNG.....	7
3	BERECHNUNGSBEISPIEL.....	9
3.1	AUFBEREITUNG DER AUSGANGSDATEN .....	9
3.2	KONZEPTION DES MODELLAUFBAUS .....	10
3.3	ERMITTLUNG DES OPTIMALEN PARAMETERVEKTORS.....	12
3.4	ERMITTLUNG DER AUSGLEICHSWERTE.....	13
4	MODELLVERGLEICHE .....	17
4.1	GEWICHTETES ADDITIVES MODELL.....	17
4.2	UNGEWICHTETES ADDITIVES MODELL.....	19
4.3	GEWICHTETES MULTIPLIKATIVES MODELL.....	21
4.4	UNGEWICHTETES MULTIPLIKATIVES MODELL .....	24
5	FAZIT .....	27
	LITERATURVERZEICHNIS .....	29
	ABBILDUNGSVERZEICHNIS .....	30

# 1 Vorbemerkungen

Die lineare Regression ist das bekannteste Verfahren zur Fehlerausgleichung, welches relativ einfach umgesetzt werden kann. Verallgemeinerte lineare Modelle sind eine zweckmäßige Erweiterung dieses Verfahrens, können aber aufgrund ihrer hohen Komplexität i. d. R. nur mit spezieller Software gerechnet werden. Aus diesem Grund wird in dieser Arbeit anhand eines fiktiven Stornobestandes illustriert, wie man Ausgleichsrechnungen auch mit Hilfe von Gauß Markow Modellen durchführen kann, die mittels EXCEL und Visual Basic noch vergleichsweise einfach umsetzbar sind.

## 1.1 Allgemeiner Ansatz einer linearen Regression

Die allgemeine Idee bei einer linearen Regression besteht darin, dass man durch eine Datenmenge die (bzgl. eines geeigneten Beurteilungskriteriums) bestmögliche Gerade legt. Dazu betrachtet man zunächst einmal die Gleichungssysteme

$$\begin{aligned} Y_1 &= a + b \cdot X_1 + \varepsilon_1 \\ Y_2 &= a + b \cdot X_2 + \varepsilon_2 \\ &\dots \\ Y_n &= a + b \cdot X_n + \varepsilon_n \end{aligned}$$

mit  ${}^tY = (Y_1, Y_2, \dots, Y_n)$  Realisationen einer Beobachtungsvariablen,  ${}^tX = (X_1, X_2, \dots, X_n)$  Realisationen einer Erklärungsvariablen und  ${}^t\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$  Realisationen einer Fehlervariablen. Gesucht werden Parameter  $(a, b)$  derart, dass die Fehlerquadrate

$$S(a, b) = \sum_i (Y_i - (a + b \cdot X_i))^2$$

minimal werden. Betrachtet man die Nullstellen der partiellen Ableitungen  $(\partial/\partial a) S$  und  $(\partial/\partial b) S$ , so erhält man die Bedingungen

$$-2 \cdot \sum_i (Y_i - (a + b \cdot X_i)) = 0 \quad \rightarrow \quad a = E[Y] - b \cdot E[X]$$

$$-2 \cdot \sum_i (Y_i - (a + b \cdot X_i)) \cdot X_i = 0 \quad \rightarrow \quad b = \text{COV}[X, Y] / \text{VAR}[X]$$

mit

$$E[Y] = 1/n \cdot \sum_i Y_i$$

$$E[X] = 1/n \cdot \sum_i X_i$$

$$\text{COV}[X, Y] = E[(X - E[X]) \cdot (Y - E[Y])]$$

$$\text{VAR}[X] = \text{COV}[X, X] = E[(X - E[X])^2].$$

Die so ermittelten Parameter  $(a, b)$  definieren eine Gerade durch die Daten derart, dass das arithmetische Mittel des Fehlers null und die Fehlervarianz minimal ist. Die Qualität der linearen Regression kann durch den Korrelationskoeffizienten

$$\rho_{X,Y} = \text{COV}[X, Y] / (\text{VAR}[X] \cdot \text{VAR}[Y])^{0.5}$$

gemessen werden. Das Bestimmtheitsmaß  $\rho^2$  gibt dabei an, wieviel Prozent der Varianz der Beobachtungswerte durch das Modell erklärt werden, wobei insbesondere gilt:

$$\text{VAR}[\varepsilon] = (1 - \rho^2) \cdot \text{VAR}[Y].$$

Die so ermittelte Gerade ist im Hinblick auf das gewählte Optimierungskriterium die bestmögliche Wahl. Das bedeutet aber nicht, dass eine Ausgleichung durch eine Gerade das beste Verfahren darstellt, da es möglicherweise andere als lineare Zusammenhänge gibt. Man muss also in jedem Fall die Struktur der Fehler analysieren und überprüfen, ob die Annahme eines linearen Zusammenhangs plausibel ist.

Es gibt mehrere relativ simple Verallgemeinerungen dieses einfachen Ansatzes, die nachfolgend kurz erläutert werden.

## 1.2 Gewichtete Modelle

In dem zuvor erläuterten Modell gehen alle Daten mit gleichem Gewicht  $1/n$  in die Fehlerausgleichung ein. Häufig hat man jedoch Gewichtungen  $(w_1, w_2, \dots, w_n)$ , d. h. manche Datenpunkte haben ein höheres Gewicht als andere, so dass eine gute Ausgleichung in diesen Punkten wichtiger ist als in anderen Punkten.

In diesem Fall ist es zweckmäßig, nicht die Fehlerquadrate  $S(a, b)$ , sondern die gewichteten Fehlerquadrate

$$S(a, b, w) = \sum_i w_i \cdot (Y_i - (a + b \cdot X_i))^2$$

zu minimieren. Man erhält konzeptionell die gleichen Lösungen – allerdings mit den arithmetischen Mitteln, Varianzen und Kovarianzen bezüglich der auf Eins normierten Gewichte.

## 1.3 Multiplikative Modelle

Betrachtet man anstelle von  $Y_1, Y_2, \dots, Y_n$  und  $X_1, X_2, \dots, X_n$  die logarithmierten Werte  $\text{LN}(Y_1), \text{LN}(Y_2), \dots, \text{LN}(Y_n)$  und  $\text{LN}(X_1), \text{LN}(X_2), \dots, \text{LN}(X_n)$ , dann ergibt sich nach Re-Transformation der linearen Regressionsgerade ein multiplikativer Zusammenhang

$$Y_i = a^* \cdot (X_i)^b \cdot \varepsilon_i^*$$

In diesem Fall ist nicht mehr die Summe des re-transformierten Fehlers  $\varepsilon^*$  null, sondern das Produkt ist gleich Eins. Somit stimmt auch nicht das arithmetische Mittel der Beobachtungen mit dem arithmetischen Mittel der modellierten Werte überein, sondern mit dem geometrischen Mittel. Bei Anwendungen von Gewichten ergibt sich hier entsprechend das gewichtete geometrische Mittel.

Logarithmiert man nur die Beobachtungen, dann erhält man einen exponentiellen Modellansatz

$$Y_i = a^* \cdot (b^*)^{X_i} \cdot \varepsilon_i^*$$

Dieser Modellansatz eignet sich gut bei Zeitreihenanalysen, wo die erklärende Variable ein Zeitparameter  $t$  ist, d. h. man hat dann mit

$$Y_t = a^* \cdot (b^*)^t \cdot \varepsilon_t^*$$

den klassischen Ansatz für multiplikative Zeittrends.

#### 1.4 Credibility Modelle

Bei manchen Zeitreihen ist der bestmögliche additive oder multiplikative Trend u. U. zu stark, so dass man – insbesondere bei sehr kleinem Bestimmtheitsmaß – für eine gute Prognose in die Zukunft nicht darauf vertrauen sollte. In diesem Fall kann man als Prognosewert den sogenannten Credibility Wert CW

$$CW = \rho^2 \cdot RW + (1 - \rho^2) \cdot EW,$$

RW der Regressionswert und EW der Erwartungswert ansetzen, wobei dieser Schätzer dann keine minimale Fehlervarianz hat. Dennoch hat der Schätzer andere vernünftige Eigenschaften:

- Er ist erwartungstreu.
- Er hat eine maximal 25% höhere Fehlervarianz als der Regressionsschätzer.
- Bei kleinem Bestimmtheitsmaß stimmt er fast mit dem Erwartungswert überein.
- Bei hohem Bestimmtheitsmaß stimmt er fast mit dem Regressionswert überein.

Dieser Schätzwert ist somit ein guter Kompromiss zwischen dem Erwartungswert und dem Regressionswert als Schätzer.

Im folgenden Abschnitt wird erläutert, wie man durch Gauß Markow Modelle weitere Verallgemeinerungen der linearen Regression erhalten kann.

## 2 Gauß Markow Modelle

In diesem Abschnitt wird zunächst der allgemeine Gauß Markow Ansatz kurz erläutert (siehe hierzu auch [1]), wobei im Anschluss illustriert wird, wie man mit Hilfe geeigneter Transformationen des Modells auch Gewichtungen mit einbeziehen kann. Dies ist gerade bei versicherungsmathematischen Analysen sehr zweckmäßig, da es hier häufig natürliche Gewichtungen durch Prämien, Jahreseinheiten, Exposures (wie etwa Versicherungssummen) oder Schadenanzahlen gibt.

### 2.1 Allgemeiner Ansatz eines Gauß Markow Modells

Die zuvor erläuterte lineare Regression kann auch als  $Y = X \cdot \beta + \varepsilon$  mit einem Beobachtungsvektor  $Y$ , einer  $(n \times m)$  Designmatrix  $X$ , einem Parametervektor  $\beta$  und einem Fehlervektor  $\varepsilon$  dargestellt werden, wobei für die Designmatrix

$$X = \begin{vmatrix} 1 & X_1 \\ \dots & \dots \\ 1 & X_n \end{vmatrix}$$

und für den Parametervektor  ${}^t\beta = (a, b)$  gilt. In einer konsequenten Matrixschreibweise gilt für die Quadrate der Fehler die Darstellung

$$\begin{aligned} S(\beta) &= {}^t(Y - X \cdot \beta) \cdot (Y - X \cdot \beta) \\ &= ({}^tY \cdot Y - 2 \cdot {}^t\beta \cdot {}^tX \cdot Y + {}^t\beta \cdot {}^tX \cdot X \cdot \beta) \end{aligned}$$

Für den Vektor der ersten partiellen Ableitungen  $(\partial/\partial\beta) S$  gilt dann die Beziehung

$$(\partial/\partial\beta) S = -2 \cdot ({}^tX \cdot Y - {}^tX \cdot X \cdot \beta) = 0,$$

woraus dann für das Minimum die Gleichung

$${}^tX \cdot X \cdot \beta = {}^tX \cdot Y$$

bzw. für den geschätzten (besten) Parametervektor

$$\beta^\circ = ({}^tX \cdot X)^{-1} \cdot ({}^tX \cdot Y)$$

folgt, sofern die Matrix  $X$  Maximalrang hat. Aus dieser Gleichung erhält man sofort eine Beziehung

$${}^tX \cdot (Y - X \cdot \beta^\circ) = {}^tX \cdot \varepsilon = 0.$$

Sofern also die Designmatrix  $X$  den Offsetvektor  ${}^tO = (1, 1, \dots, 1)$  als Spalte enthält (was bei der linearen Regression der Fall ist), dann ergibt sich für die Summe der Abweichungen zwischen den Beobachtungen und den modellierten Werten die Beziehung

$$\sum_i 1 \cdot (Y_i - (\beta^\circ_1 \cdot 1 + \beta^\circ_2 \cdot X_{i,2} + \dots + \beta^\circ_m \cdot X_{i,m})) = \sum_i \varepsilon_i = 0,$$

d. h. die modellierten Werte sind erwartungstreu und die Bedingung der kleinsten Quadrate korrespondiert zu der Bedingung einer minimalen Fehlervarianz. Unter der Annahme eines normalverteilten Fehlers (z. B. bei Anwendung des zentralen Grenzwertsatzes) ist der so geschätzte Parametervektor  $\beta^\circ$  dann auch ein Maximum Likelihood Schätzer. Bei anderen Verteilungsannahmen gilt das nicht unbedingt; hier muss man ggf. mit anderen Methoden wie beispielsweise verallgemeinerten linearen Modellen arbeiten.

Die zuvor skizzierte Designmatrix für die lineare Regression kann jetzt beispielsweise sehr leicht um weitere Terme wie folgt ergänzt werden:

$$X = \begin{pmatrix} 1 & X_1 & X_1^2 & \dots & X_1^k \\ \dots & \dots & \dots & \dots & \dots \\ 1 & X_n & X_n^2 & \dots & X_n^k \end{pmatrix}$$

Eine Gauß Markow Ausgleichung liefert hier das bestmögliche Polynom k-ten Grades. Durch Logarithmieren kann man auch multiplikative Zusammenhänge modellieren – etwa bei Annahme von lognormalverteilten Zufallsvariablen. Allerdings minimiert man dann den Fehler nicht in der normalen Darstellung, sondern nur in der transformierten Darstellung. In diesem Fall ergibt sich die Erwartungstreu nur im Hinblick auf das geometrische Mittel.

## 2.2 Gewichtete Modelle

Für die Optimierung unter Berücksichtigung einer Gewichtung kann man zunächst einmal das transformierte Modell

$$V \cdot Y = V \cdot X \cdot \beta + V \cdot \varepsilon,$$

mit einer beliebigen Transformationsmatrix  $V$  betrachten. Sofern der transformierte Fehler

$${}^t(V \cdot \varepsilon) \cdot (V \cdot \varepsilon) = {}^t\varepsilon \cdot ({}^tV \cdot V) \cdot \varepsilon$$

minimiert werden soll, ergibt sich eine Lösung

$$\beta^\circ = ({}^tX^* \cdot X^*)^{-1} \cdot ({}^tX^* \cdot Y^*)$$

mit  $X^* = V \cdot X$  und  $Y^* = V \cdot Y$ . Falls  $V$  eine Diagonalmatrix ist, gilt  ${}^tV = V$  bzw.  ${}^tV \cdot V = V^2$ . Bei einer Gewichtung  $w_1, \dots, w_n$  ergibt sich dann für

$$W = \text{Diag}(\dots w_i \dots) = \text{Diag}(\dots w_i^{0.5} \dots)^2 = V^2$$

für die gewichteten Fehlerquadrate

$${}^t\varepsilon \cdot W \cdot \varepsilon = {}^t\varepsilon \cdot V^2 \cdot \varepsilon = {}^t\varepsilon \cdot ({}^tV \cdot V) \cdot \varepsilon$$

ein modifizierter Optimierungsansatz

$$\beta^\circ = ({}^tX \cdot W \cdot X)^{-1} \cdot ({}^tX \cdot W \cdot Y) = ({}^tX^* \cdot X^*)^{-1} \cdot ({}^tX^* \cdot Y^*).$$

Bei diesem Modell werden die quadratischen Abweichungen des transformierten Fehlervektors  ${}^t\epsilon^* = (v_1 \cdot \epsilon_1, \dots, v_n \cdot \epsilon_n)$  minimiert. Sofern die Designmatrix  $X$  einen Offsetvektor enthält, wird dieser in der transformierten Designmatrix  $X^*$  zu einem Vektor  ${}^t(v_1, \dots, v_n)$ , so dass die Summe der Fehler jetzt nicht mehr null ist. Für die Fehlerabweichungen gelten die Beziehungen

$$\begin{aligned} \sum_i v_i \cdot \epsilon_i^* &= \sum_i v_i^2 \cdot \epsilon_i &= \sum_i w_i \cdot \epsilon_i &= 0 \\ \sum_i (\epsilon_i^*)^2 &= \sum_i v_i^2 \cdot \epsilon_i^2 &= \sum_i w_i \cdot \epsilon_i^2 &= \text{minimal.} \end{aligned}$$

Sofern die Gewichte auf Eins normiert sind, sind die modellierten Werte also im Hinblick auf den gewichteten Erwartungswert erwartungstreu bei minimaler (gewichteter) Fehlervarianz.

### 2.3 Qualitative Merkmale

Bei qualitativen Merkmalen ergibt sich ein anderer Aufbau für die Designmatrix. Hierzu betrachtet man  $m_{i,j}$  definiert als die  $j$ -te Ausprägung des  $i$ -ten Merkmals. Setzt man eine Offsetvariable für eine Basisausprägung für jedes Merkmal  $m_i$ , dann ist die Designmatrix wie folgt definiert:

$$X = \begin{vmatrix} 1 & \dots & \dots & \dots \\ \dots & \dots & \delta_{i,j,k} & \dots \\ 1 & \dots & \dots & \dots \end{vmatrix}$$

mit  $\delta_{i,j,k} = (m_i(k) = m_{i,i}) \cdot 1$  für  $m_i(k)$  die Ausprägung des  $i$ -ten Merkmals des Beobachtungswerts  $Y_k$  (mit Ausnahme der Basisausprägungen, die in der Offsetvariablen abgebildet sind). Der Gauß Markow Ansatz entspricht jetzt einem Gleichungssystem

$$Y_k = \beta_0 + \sum \beta_{i,j} \cdot \delta_{i,j,k} + \epsilon_k.$$

Bei einem multiplikativen Modell ergibt sich das Gleichungssystem

$$\text{LN}(Y_k) = \beta_0 + \sum \beta_{i,j} \cdot \delta_{i,j,k} + \epsilon_k \quad \text{bzw.}$$

$$Y_k = f_0 \cdot \prod f_{i,j,k} \cdot E_k$$

mit  $f_{i,j,k} = 1$  falls  $m_i(k) \neq m_{i,i}$  und  $E_k = \text{EXP}(\epsilon_k)$ . Dieser Ansatz soll in dieser Arbeit für die Stornowahrscheinlichkeiten

$$Y_k = \text{ST}_k / \text{JE}_{\text{max},k}$$

für  $\text{ST}_k$  die Anzahl der Stornos und  $\text{JE}_{\text{max},k}$  die (maximalen) Jahreseinheiten in der  $k$ -ten Zelle diskutiert werden. Hierbei ist zu beachten, dass man für eine korrekte Einschätzung der Stornowahrscheinlichkeiten nicht die Jahreseinheiten bis zum Stornozeitpunkt, sondern die im konkreten Jahr maximal möglichen Jahreseinheiten ansetzt. (Für eine präzise Herleitung siehe [2].)

Damit dieser Modellansatz funktioniert, sollten für alle denkbaren Merkmalskombinationen auch Jahreseinheiten vorhanden sein. Dies erzielt man beispielsweise, indem man einen großen Bestand wählt oder die Anzahl der Merkmalsausprägungen reduziert.

### 2.3.1 Additiver Modellansatz

Sofern die Jahreseinheiten von Null verschieden sind, funktioniert dieser Ansatz in jedem Fall. Selbst wenn man nicht für alle Merkmalskombinationen Beobachtungen hat, kann man diese durch „Randwerte“ approximieren.

Aus  $w_k = JE_{\max, k} / JE_{\max, ges}$  ergibt sich eine natürliche Gewichtung. Dies korrespondiert zur „Optimierung“ der gesamten Stornowahrscheinlichkeit

$$SW_{ges} = \sum w_k \cdot SW_k$$

mit normalverteilten Fehlern. Es gibt jedoch ein Problem: Aufgrund des linearen Ansatzes können sich auch negative Werte für die Stornowahrscheinlichkeiten aus diesem Optimierungsansatz ergeben. Dies kann durch einen multiplikativen Ansatz behoben werden.

### 2.3.2 Multiplikativer Modellansatz

Wie bereits erläutert korrespondiert dieser Ansatz zur Optimierung des (gewichteten) geometrischen Mittels bei der Annahme von lognormalverteilten Variablen. Hier ergibt sich das Problem, dass bei beobachteter Stornowahrscheinlichkeit von null der Logarithmus nicht definiert ist. Eliminiert man diese Werte aus der Optimierung, können diese aus den Randwerten heraus abgeleitet werden. Allerdings ergibt sich dann das Problem eines „Overfits“, d. h. alle Werte ungleich null werden gut getroffen und für Werte gleich null werden positive Werte geschätzt (was Sinn macht). Insgesamt ist die Stornowahrscheinlichkeit also zu hoch eingeschätzt.

Hier ergibt somit die Besonderheit, dass man den Bestand vorher derart gut definieren muss, dass in jeder Zelle auch Storno beobachtet wird. Sofern dies nicht der Fall ist, muss man ggf. mit einer sehr kleinen Zahl sowohl die Gewichtungen als auch die Stornowahrscheinlichkeiten vorbelegen.

## 2.4 Technische Umsetzung mit Hilfe der Cholesky Zerlegung

Da die Matrix  $A = 'X \cdot X$  symmetrisch ist, kann man hier für die Lösung des Gleichungssystems

$$A \cdot x = y$$

mit dem Cholesky Verfahren arbeiten, das programmtechnisch (z. B. in Visual Basic) relativ leicht umzusetzen ist. Dieses Verfahren hat gegenüber anderen Verfahren (z. B. der EXCEL Matrix Invertierung) den Vorteil, dass es numerisch sehr stabil ist. Beim Cholesky Verfahren wird die Matrix A zerlegt in

$$A = C \cdot 'C$$

mit  $C$  eine untere Dreiecksmatrix, deren Einträge sukzessive ermittelt werden können. So gilt beispielsweise für die erste Spalte der Matrix  $C$ :

$$\begin{aligned} a_{1,1} &= c_{1,1} \cdot c_{1,1} && \rightarrow && c_{1,1} &= a_{1,1}^{0,5} \\ a_{2,1} &= c_{2,1} \cdot c_{1,1} && \rightarrow && c_{2,1} &= a_{2,1} / c_{1,1} \\ \dots &&& \dots && \dots & \\ a_{n,1} &= c_{n,1} \cdot c_{1,1} && \rightarrow && c_{n,1} &= a_{n,1} / c_{1,1} \end{aligned}$$

Für die zweite Spalte von  $C$  gilt dann:

$$\begin{aligned} a_{2,2} &= c_{2,1} \cdot c_{2,1} + c_{2,2} \cdot c_{2,2} && \rightarrow && c_{2,2} &= (a_{2,2} - c_{2,1} \cdot c_{2,1})^{0,5} \\ a_{3,2} &= c_{2,1} \cdot c_{3,1} + c_{2,2} \cdot c_{3,2} && \rightarrow && c_{3,2} &= (a_{3,2} - c_{2,1} \cdot c_{3,1}) / c_{2,2} \\ \dots &&& \dots && \dots & \\ a_{n,2} &= c_{2,1} \cdot c_{n,1} + c_{2,2} \cdot c_{n,2} && \rightarrow && c_{n,2} &= (a_{n,2} - c_{2,1} \cdot c_{n,1}) / c_{2,2} \end{aligned}$$

Das Verfahren geht jetzt sukzessive so weiter, wobei der Algorithmus relativ einfach in einer Programmierschleife umsetzbar ist. Man kann dann aufgrund der Eigenschaften als unterer bzw. oberer Diagonalmatrix recht einfach die Gleichungssysteme  $C \cdot z = y$  und  ${}^t C \cdot x = z$  lösen.

### 3 Berechnungsbeispiel

Im Folgenden wird für das Beispiel von Stornowahrscheinlichkeiten ein Gauß Markow Modell in EXCEL durchgerechnet.

#### 3.1 Aufbereitung der Ausgangsdaten

Als Ausgangsbasis liegt dabei ein komprimierter (virtuell erzeugter) Bestand mit folgender Datenstruktur vor (siehe [2]):

Datenfeld	Ausprägungen	Erläuterungen
Jahr	2009, 2010, 2011, 2012, 2013	Beobachtungsjahr
AG	18 (= 18 – 24), 25 (= 25 – 34), 35 (= 35 – 54), 55 (= 55 – 74), 75 (= ab 75)	Altersgruppe im Beobachtungsjahr
Geschl.	0 = Männlich, 1 = Weiblich	Geschlecht
Verh.	0 = Nein, 1 = Ja	Familienstand
Beamter	0 = Nein, 1 = Ja	Berufsgruppe
High Risk	0 = Nein, 1 = Ja	Risikogruppe
PG	0% (= niedrig, bis 95%) 95% (= mittel, 95% bis 105%) 105% (= hoch, ab 105%)	Profitabilitätsgruppe definiert durch den Prozentsatz $P = \text{Ist-Prämie} / \text{Soll-Prämie}$

In der nachfolgenden Abbildung ist ein Teil der modellierten Daten zusammen mit der Summe der Jahreseinheiten und der durchschnittlichen Stornowahrscheinlichkeit aufgelistet:

Jahr	Alter	Geschl.	Verh.	Beamter	High Risk	PG	STW	JE
							0,067	31.375,3
2013	18	0	0	0	0	0%	0,119	58,8
2013	18	0	0	0	0	95%	0,000	7,0
2013	18	0	0	0	0	105%	0,094	42,7
2013	18	0	0	0	1	0%	0,135	7,4
2013	18	0	0	0	1	95%	0,286	3,5
2013	18	0	0	0	1	105%	0,000	8,0
2013	18	0	0	1	0	0%	0,050	19,9
2013	18	0	0	1	0	95%	0,000	2,0
2013	18	0	0	1	0	105%	0,000	10,0
2013	18	0	0	1	1	0%	0,000	2,0
2013	18	0	0	1	1	95%	0,000	1,0
2013	18	0	0	1	1	105%	0,211	4,8
2013	18	0	1	0	0	0%	0,064	78,0
2013	18	0	1	0	0	95%	0,000	11,0
2013	18	0	1	0	0	105%	0,034	58,3
2013	18	0	1	0	1	0%	0,058	17,2
2013	18	0	1	0	1	95%	0,316	3,2
2013	18	0	1	0	1	105%	0,000	6,0
2013	18	0	1	1	0	0%	0,083	36,1
2013	18	0	1	1	0	95%	0,000	1,0
2013	18	0	1	1	0	105%	0,153	19,7
2013	18	0	1	1	1	0%	0,138	7,3
2013	18	0	1	1	1	95%	0,000	0,0
2013	18	0	1	1	1	105%	0,000	1,0

**Abbildung 1: Auszug aus der komprimierten Beobachtungsdatei.**

Die beobachtete Stornowahrscheinlichkeit soll ausgeglichen werden, wobei bei  $STW = 0$  ein sehr kleiner Wert vorbelegt wurde, damit auch ein multiplikatives Modell gerechnet werden kann. Darüber hinaus wurden die fiktiv modellierten Stornowahrscheinlichkeiten auf 100% begrenzt.

Analog wurde bei  $JE = JE_{max} = 0$  ebenfalls ein sehr kleiner Wert vorbelegt, damit alle Kombinationen in das Modell einbezogen werden können. Da die Stornowahrscheinlichkeit für jede Merkmalskombination als

$$STW = \text{Anzahl Storno} / \text{Jahreseinheiten}$$

definiert ist, definieren die (maximalen) Jahreseinheiten auf natürliche Art und Weise eine Gewichtung. Je mehr Jahreseinheiten bei einer Merkmalskombination  $m$  vorhanden sind, desto glaubwürdiger ist die beobachtete Stornowahrscheinlichkeit. Aus diesem Grund sollten optimaler Weise die gewichteten Fehlerquadrate  ${}^t \epsilon \cdot W \cdot \epsilon$  mit  $W = \text{Diag}(\dots JE_{max, k} \dots)$  minimiert werden.

### 3.2 Konzeption des Modellaufbaus

Wie zuvor erläutert, ergibt sich die Lösung dieses Ansatzes, indem man sowohl die Designmatrix als auch den Beobachtungsvektor mit  $V = \text{Diag}(\dots JE_{max, k}^{0,5} \dots)$  transformiert. Für die Lösung dieses Modells ist dann nicht die Fehlersumme, sondern die mit  $V$  gewichtete Fehlersumme null. In der nachfolgenden Abbildung ist ein Auszug aus dem transformierten Beobachtungsvektor aufgelistet:

Jahr	Alter	Geschl.	Verh.	Beamter	High Risk	PG	STW mod	JE mod	V	W	Y*
							0,067	31.375,4	0,5	31.375,4	0,067
2013	18	0	0	0	0	0%	0,119	58,8	7,7	58,8	0,913
2013	18	0	0	0	0	95%	0,000	7,0	2,6	7,0	0,000
2013	18	0	0	0	0	105%	0,094	42,7	6,5	42,7	0,612
2013	18	0	0	0	1	0%	0,135	7,4	2,7	7,4	0,367
2013	18	0	0	0	1	95%	0,286	3,5	1,9	3,5	0,535
2013	18	0	0	0	1	105%	0,000	8,0	2,8	8,0	0,000
2013	18	0	0	1	0	0%	0,050	19,9	4,5	19,9	0,224
2013	18	0	0	1	0	95%	0,000	2,0	1,4	2,0	0,000
2013	18	0	0	1	0	105%	0,000	10,0	3,2	10,0	0,000
2013	18	0	0	1	1	0%	0,000	2,0	1,4	2,0	0,000
2013	18	0	0	1	1	95%	0,000	1,0	1,0	1,0	0,000
2013	18	0	0	1	1	105%	0,211	4,8	2,2	4,8	0,459
2013	18	0	1	0	0	0%	0,064	78,0	8,8	78,0	0,566
2013	18	0	1	0	0	95%	0,000	11,0	3,3	11,0	0,000
2013	18	0	1	0	0	105%	0,034	58,3	7,6	58,3	0,262
2013	18	0	1	0	1	0%	0,058	17,2	4,1	17,2	0,241
2013	18	0	1	0	1	95%	0,316	3,2	1,8	3,2	0,562
2013	18	0	1	0	1	105%	0,000	6,0	2,4	6,0	0,000
2013	18	0	1	1	0	0%	0,083	36,1	6,0	36,1	0,499
2013	18	0	1	1	0	95%	0,000	1,0	1,0	1,0	0,000
2013	18	0	1	1	0	105%	0,153	19,7	4,4	19,7	0,676
2013	18	0	1	1	1	0%	0,138	7,3	2,7	7,3	0,371
2013	18	0	1	1	1	95%	0,000	0,0	0,0	0,0	0,000
2013	18	0	1	1	1	105%	0,000	1,0	1,0	1,0	0,000

**Abbildung 2: Auszug aus den transformierten Werten – gewichtetes additives Modell.**

Wie zuvor erläutert muss beim Ansatz der Designmatrix eine Basis Merkmalskombination ausgewählt werden, die in der Offsetvariablen abgebildet wird (sonst hat die Matrix X nicht maximalen Rang und  $X \cdot X$  ist nicht invertierbar). Im vorliegenden Beispiel ist dies die Kombination

(Jahr = 2013, Alter = 18, Geschl. = 0, Verh. = 0, Beamter = 0, High Risk = 0, PG = 0%).

Diese Kombination taucht dann in der Designmatrix nicht mehr explizit auf. In der nachfolgenden Abbildung ist ein Auszug aus der mit V transformierten Designmatrix  $X^*$  abgebildet; diese unterscheidet sich nur bei einem gewichteten oder ungewichteten Modell, nicht aber bei einem additiven oder multiplikativen Modell.

X*	Jahr				Alter				G	VH	B	HR	PG		
	Offset	2012	2011	2010	2009	25	35	55	75	1	1	1	1	95%	105%
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
3	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0
3	0	0	0	0	0	0	0	0	0	0	0	3	0	0	3
4	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
3	0	0	0	0	0	0	0	0	0	0	3	0	0	0	3
1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
2	0	0	0	0	0	0	0	0	0	0	2	2	0	0	2
9	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	3	0	0	3	0	0
8	0	0	0	0	0	0	0	0	0	8	0	0	0	0	8
4	0	0	0	0	0	0	0	0	0	4	0	4	0	0	0
2	0	0	0	0	0	0	0	0	0	2	0	2	2	0	0
2	0	0	0	0	0	0	0	0	0	2	0	2	0	0	2
6	0	0	0	0	0	0	0	0	0	6	6	0	0	0	0
1	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0
4	0	0	0	0	0	0	0	0	0	4	4	0	0	0	4
3	0	0	0	0	0	0	0	0	0	3	3	3	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1

Abbildung 3: Auszug aus der transformierten Designmatrix – gewichtetes Modell.

Sofern man mit einem ungewichteten Modell arbeitet, steht anstelle der unterschiedlichen Gewichte überall der Wert Eins.

### 3.3 Ermittlung des optimalen Parametervektors

Die Matrix  ${}^tX^* \cdot X^*$  ist in der nachfolgenden Abbildung dargestellt:

	Offset	2012	2011	2010	2009	25	35	55	75	1	1	1	1	95%	105%
Offset	31375	5968	6196	6688	6876	8750	11688	4179	1433	12173	19167	7988	4940	2535	11498
2012	5968	5968				1658	2327	828	314	2342	3613	1527	887	482	2192
2011	6196		6196			1764	2322	834	301	2443	3745	1583	925	499	2265
2010	6688			6688		1856	2371	835	267	2468	4232	1688	1274	530	2468
2009	6876				6876	1918	2384	857	236	2699	4167	1738	1018	559	2513
25	8750	1658	1764	1856	1918	8750				3438	5295	2292	1277	658	3164
35	11688	2327	2322	2371	2384		11688			4473	7143	2924	1758	923	4266
55	4179	828	834	835	857			4179		1695	2632	1032	577	401	1520
75	1433	314	301	267	236				1433	614	795	378	226	129	529
1	12173	2342	2443	2468	2699	3438	4473	1695	614	12173	7480	3173	1869	1037	4444
1	19167	3613	3745	4232	4167	5295	7143	2632	795	7480	19167	4826	3219	1506	7047
1	7988	1527	1583	1688	1738	2292	2924	1032	378	3173	4826	7988	1283	667	2807
1	4940	887	925	1274	1018	1277	1758	577	226	1869	3219	1283	4940	410	1798
95%	2535	482	499	530	559	658	923	401	129	1037	1506	667	410	2535	
105%	11498	2192	2265	2468	2513	3164	4266	1520	529	4444	7047	2807	1798		11498

Abbildung 4: Transformierte Lösungsmatrix – gewichtetes additives Modell.

Der erste Teil der der Cholesky Zerlegung  ${}^tX^* \cdot X^* = C \cdot {}^tC$  in die untere Dreiecksmatrix C zusammen mit der Lösung des Gleichungssystems  $C \cdot z = {}^tX^* \cdot Y^*$  ist nachfolgend abgebildet:

Ziel	Lösung	Offset	2012	2011	2010	2009	25	35	55	75	1	1	1	1	95%	105%
$\text{X}^* \cdot \text{Y}^*$	Hilfsw.	11,931	0,583	0,405	-0,007	-2,201	0,332	0,246	0,104	0,316	0,070	0,364	-0,279	-0,103	-0,155	0,078
2113,372	11,931	177														
442,507	0,583	34	70													
435,174	0,405	35	-17	68												
429,675	-0,007	38	-18	-24	66											
320,010	-2,201	39	-19	-24	-36	56										
615,425	0,332	49	0	1	0	0	79									
800,341	0,246	66	1	1	-1	-3	-41	75								
277,842	0,104	24	0	0	-1	-1	-15	-29	51							
105,006	0,316	8	1	0	0	-1	-5	-10	-11	34						
827,352	0,070	69	0	1	-2	0	1	-1	1	2	86					
1320,853	0,364	108	0	-1	2	0	-1	0	1	-2	1	86				
517,187	-0,279	45	0	0	0	0	1	0	-1	0	1	-1	77			
324,353	-0,103	28	-1	-1	3	0	-1	-2	-3	-2	0	2	0	64		
163,011	-0,155	14	0	0	0	0	-1	-1	1	0	1	-1	0	0	48	
784,015	0,078	65	0	0	0	0	-1	-1	-1	0	0	0	-2	0	-19	83

Abbildung 5: Cholesky Zerlegung – gewichtetes additives Modell (1).

Der zweite Teil der der Cholesky Zerlegung  $\text{X}^* \cdot \text{X}^* = \text{C} \cdot \text{C}$  in die obere Dreiecksmatrix  $\text{C}$  zusammen mit der Lösung des Gleichungssystems  $\text{C} \cdot \beta^* = z$  ist nachfolgend abgebildet:

Ziel	Lösung	Offset	2012	2011	2010	2009	25	35	55	75	1	1	1	1	95%	105%
Hilfsw.	Beta	0,078	-0,012	-0,016	-0,021	-0,039	0,009	0,006	0,004	0,010	0,001	0,004	-0,004	-0,002	-0,003	0,001
11,931	0,078	177	34	35	38	39	49	66	24	8	69	108	45	28	14	65
0,583	-0,012		70	-17	-18	-19	0	1	0	1	0	0	0	-1	0	0
0,405	-0,016			68	-24	-24	1	1	0	0	1	-1	0	-1	0	0
-0,007	-0,021				66	-36	0	-1	-1	0	-2	2	0	3	0	0
-2,201	-0,039					56	0	-3	-1	-1	0	0	0	0	0	0
0,332	0,009						79	-41	-15	-5	1	-1	1	-1	-1	-1
0,246	0,006							75	-29	-10	-1	0	0	-2	-1	-1
0,104	0,004								51	-11	1	1	-1	-3	1	-1
0,316	0,010									34	2	-2	0	-2	0	0
0,070	0,001										86	1	1	0	1	0
0,364	0,004											86	-1	2	-1	0
-0,279	-0,004												77	0	0	-2
-0,103	-0,002													64	0	0
-0,155	-0,003														48	-19
0,078	0,001															83

Abbildung 6: Cholesky Zerlegung – gewichtetes additives Modell (2).

Die hier dargestellte Cholesky Zerlegung wurde explizit in EXCEL auf Basis einer fest vorgegebenen Struktur gerechnet. Das Verfahren kann aber relativ einfach mit Visual Basic dynamisiert werden.

### 3.4 Ermittlung der Ausgleichswerte

Durch Anwendung von  $\beta^\circ$  auf  $\text{X}^*$  erhält man jetzt die modellierten Werte für den transformierten Beobachtungsvektor  $\text{Y}^*$ , die man dann wieder auf die Ausgangswerte zurück transformieren kann. Wie bereits erläutert ist der gewichtete Fehlererwartungswert für die transformierten Werte null, siehe dazu auch den nachfolgenden Auszug aus den Ergebnissen.

Jahr	Alter	Geschl.	Verh.	Beamter	Y*	X* · β°	ε*	STW°	Abw.
					0,067	0,000	0,003	0,000	0,003
2013	18	0	0	0	0,913	0,601	0,312	0,078	0,041
2013	18	0	0	0	0,000	0,200	-0,200	0,075	-0,075
2013	18	0	0	0	0,612	0,518	0,095	0,079	0,014
2013	18	0	0	0	0,367	0,209	0,158	0,077	0,058
2013	18	0	0	0	0,535	0,138	0,396	0,074	0,212
2013	18	0	0	0	0,000	0,220	-0,220	0,078	-0,078
2013	18	0	0	1	0,224	0,333	-0,109	0,075	-0,025
2013	18	0	0	1	0,000	0,102	-0,102	0,072	-0,072
2013	18	0	0	1	0,000	0,239	-0,239	0,076	-0,076
2013	18	0	0	1	0,000	0,103	-0,103	0,073	-0,073
2013	18	0	0	1	0,000	0,070	-0,070	0,070	-0,070
2013	18	0	0	1	0,459	0,161	0,297	0,074	0,136
2013	18	0	1	0	0,566	0,729	-0,163	0,083	-0,018
2013	18	0	1	0	0,000	0,264	-0,264	0,080	-0,080
2013	18	0	1	0	0,262	0,638	-0,376	0,083	-0,049
2013	18	0	1	0	0,241	0,335	-0,094	0,081	-0,023
2013	18	0	1	0	0,562	0,139	0,423	0,078	0,238
2013	18	0	1	0	0,000	0,201	-0,201	0,082	-0,082
2013	18	0	1	1	0,499	0,474	0,025	0,079	0,004
2013	18	0	1	1	0,000	0,076	-0,076	0,076	-0,076
2013	18	0	1	1	0,676	0,354	0,322	0,080	0,073
2013	18	0	1	1	0,371	0,208	0,163	0,077	0,061
2013	18	0	1	1	0,000	0,002	-0,002	0,075	-0,075
2013	18	0	1	1	0,000	0,078	-0,078	0,078	-0,078

**Abbildung 7: Auszug aus den modellierten Werten – gewichtetes additives Modell.**

Alternativ wurde auch ein gewichtetes multiplikatives Modell gerechnet, wobei sich dabei nur der Beobachtungsvektor geändert hat; Gewichtung und Designmatrix blieben gleich. In der nachfolgenden Abbildung sind die Ergebnisse (nach Re-Transformation durch die Exponentialfunktion) hierfür aufgelistet.

Jahr	Alter	Geschl.	Verh.	Beamter	Y*	X* · β°	ε*	STW°	Abw.
					-3,807	0,000	6,208	0,019	0,004
2013	18	0	0	0	-16,328	-21,063	4,735	0,064	0,055
2013	18	0	0	0	-30,460	-14,378	-16,083	0,004	-0,004
2013	18	0	0	0	-15,462	-20,671	5,209	0,042	0,052
2013	18	0	0	0	-5,457	-14,850	9,393	0,004	0,131
2013	18	0	0	0	-2,344	-15,230	12,887	0,000	0,285
2013	18	0	0	0	-32,563	-16,606	-15,957	0,003	-0,003
2013	18	0	0	1	-13,351	-18,652	5,301	0,015	0,035
2013	18	0	0	1	-16,282	-9,712	-6,570	0,001	-0,001
2013	18	0	0	1	-36,407	-14,540	-21,868	0,010	-0,010
2013	18	0	0	1	-16,282	-9,738	-6,543	0,001	-0,001
2013	18	0	0	1	-11,513	-9,574	-1,939	0,000	0,000
2013	18	0	0	1	-3,396	-15,920	12,524	0,001	0,210
2013	18	0	1	0	-24,263	-19,115	-5,148	0,115	-0,051
2013	18	0	1	0	-38,184	-16,094	-22,090	0,008	-0,008
2013	18	0	1	0	-25,762	-19,726	-6,036	0,076	-0,041
2013	18	0	1	0	-11,779	-20,182	8,403	0,008	0,051
2013	18	0	1	0	-2,051	-13,452	11,401	0,001	0,315
2013	18	0	1	0	-28,201	-12,956	-15,244	0,005	-0,005
2013	18	0	1	1	-14,941	-21,611	6,670	0,027	0,056
2013	18	0	1	1	-11,513	-6,286	-5,227	0,002	-0,002
2013	18	0	1	1	-8,339	-17,810	9,471	0,018	0,135
2013	18	0	1	1	-5,334	-16,975	11,641	0,002	0,136
2013	18	0	1	1	-0,364	-0,284	-0,080	0,000	0,000
2013	18	0	1	1	-11,513	-6,723	-4,790	0,001	-0,001

**Abbildung 8: Auszug aus den modellierten Werten – gewichtetes multiplikatives Modell.**

Zwar ist auch beim gewichteten multiplikativen Modell die gewichtete Fehlersumme gleich null und die gewichtete Fehlervarianz minimal. Da durch die Logarithmierung aber auch sehr weit nach unten gehende negative Werte auftreten, ist die Fehlervarianz deutlich höher als beim gewichteten additiven Modell.

Beim multiplikativen Modell ist wie bereits erläutert nach der Re-Transformation der gewichtete Erwartungswert der Abweichungen nicht mehr null. Im vorliegenden Beispiel ist aufgrund der sehr hohen Fehlervarianz bei den Logarithmen nach Anwendung der Exponentialfunktion die Abweichung sogar erheblich, was auf eine eher „bescheidene“ Modelleignung hindeutet.

Sofern man das multiplikative Modell dennoch wählen möchte, muss man die modellierten Stornowahrscheinlichkeiten anpassen, damit man zumindest Erwartungstreue erhält, siehe dazu die nachfolgende Abbildung.

Jahr	Alter	Geschl.	Verh.	Beamter	STW°	Abw.	STW° mod	Abw. mod
					<b>0,019</b>	<b>0,004</b>	<b>27,7%</b>	<b>0,006</b>
2013	18	0	0	0	0,064	0,055	0,082	0,037
2013	18	0	0	0	0,004	-0,004	0,006	-0,006
2013	18	0	0	0	0,042	0,052	0,054	0,040
2013	18	0	0	0	0,004	0,131	0,005	0,129
2013	18	0	0	0	0,000	0,285	0,000	0,285
2013	18	0	0	0	0,003	-0,003	0,004	-0,004
2013	18	0	0	1	0,015	0,035	0,020	0,031
2013	18	0	0	1	0,001	-0,001	0,001	-0,001
2013	18	0	0	1	0,010	-0,010	0,013	-0,013
2013	18	0	0	1	0,001	-0,001	0,001	-0,001
2013	18	0	0	1	0,000	0,000	0,000	0,000
2013	18	0	0	1	0,001	0,210	0,001	0,210
2013	18	0	1	0	0,115	-0,051	0,147	-0,083
2013	18	0	1	0	0,008	-0,008	0,010	-0,010
2013	18	0	1	0	0,076	-0,041	0,096	-0,062
2013	18	0	1	0	0,008	0,051	0,010	0,048
2013	18	0	1	0	0,001	0,315	0,001	0,315
2013	18	0	1	0	0,005	-0,005	0,006	-0,006
2013	18	0	1	1	0,027	0,056	0,035	0,048
2013	18	0	1	1	0,002	-0,002	0,002	-0,002
2013	18	0	1	1	0,018	0,135	0,023	0,130
2013	18	0	1	1	0,002	0,136	0,002	0,136
2013	18	0	1	1	0,000	0,000	0,000	0,000
2013	18	0	1	1	0,001	-0,001	0,002	-0,002

**Abbildung 9: Auszug aus den angepassten Werten – gewichtetes multiplikatives Modell.**

Die Modifizierung der Modellwerte liefert erwartungstreue Werte mit vergleichsweise geringer Fehlervarianz, wobei hier aber keine Interpretation als Maximum Likelihoodschätzer möglich ist. In dieser Hinsicht passt das Gauß Markow Modell nur wirklich gut bei einem additiven Modell.

## 4 Modellvergleiche

Im Abschnitt zuvor wurden die Berechnungen für ein gewichtetes additives Gauß Markow Modell von Stornowahrscheinlichkeiten erläutert und ein erster Vergleich mit einem gewichteten multiplikativen Modell vorgenommen. In diesem Abschnitt wird ergänzend dazu eine intensive Diskussion der Modelleigenschaft durchgeführt, insbesondere für das

- additive bzw. multiplikative Modell und das
- gewichtete bzw. ungewichtete Modell.

Dabei werden Kriterien zur Beurteilung der Modellqualität eingeführt und diskutiert.

### 4.1 Gewichtetes additives Modell

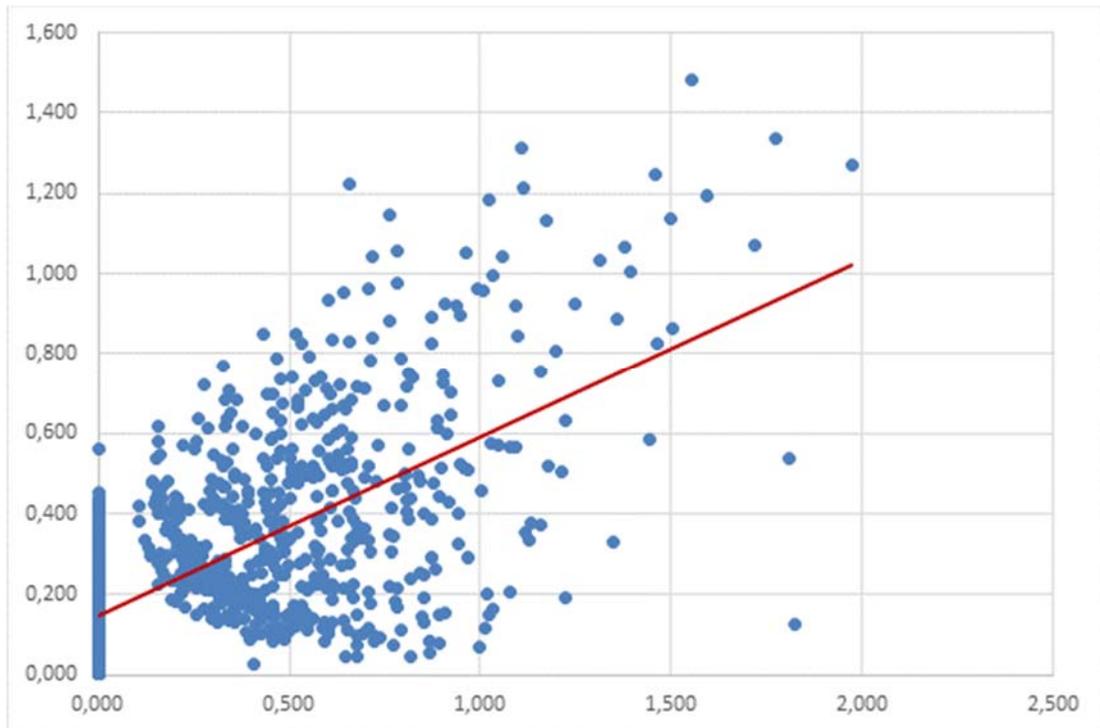
Das gewichtete additive Modell ist im Hinblick auf die gewählte Gewichtung erwartungstreu, so dass keine Modifikation in dieser Hinsicht nötig ist. Die wichtigsten Resultate dieses Modells sind nachfolgend zusammengefasst:

Gewichteter Fehlererwartungswert	0,000
Gewichtete Fehlervarianz	0,003
Abweichung Erwartungstreue in %	0,0%
Minimum modellierte Werte	3,1%
Maximum modellierte Werte	9,4%
Fehlervarianz modellierte Werte	0,003
Korrelation transformierte Werte	66,4%
Korrelation Ausgangswerte	10,0%
Korrelation Ausgangswerte gewichtet	26,3%

Die gewichtete Fehlervarianz ist vergleichsweise klein, wobei beim additiven Modell kein Unterschied zwischen der Fehlervarianz des transformierten und des Beobachtungswertes vorliegt.

Die modellierten Werte für die Stornowahrscheinlichkeiten liegen innerhalb einer sehr kleinen Bandbreite. Bei den transformierten Werten (für die die Ausgleichung ja durchgeführt wurde) ist die Korrelation zwischen transformierten Beobachtungen und den ausgeglichenen Werten recht hoch; die Korrelation zwischen den Beobachtungswerten und den zurück transformierten ausgeglichenen Werten ist mit 26,3% aber nicht besonders ausgeprägt.

In der nachfolgenden Abbildung sind die beobachteten Werte gegen die modellierten Werte für die **transformierten** Werte geplottet (wobei im Idealfall alle Werte auf der Gerade  $Y = X$  liegen sollten). Darüber hinaus ist die zu den Daten gehörende Regressionsgerade in die geplotteten Daten eingefügt worden.



**Abbildung 10: Beobachtung vs. Modell – gewichtetes additives Modell (1).**

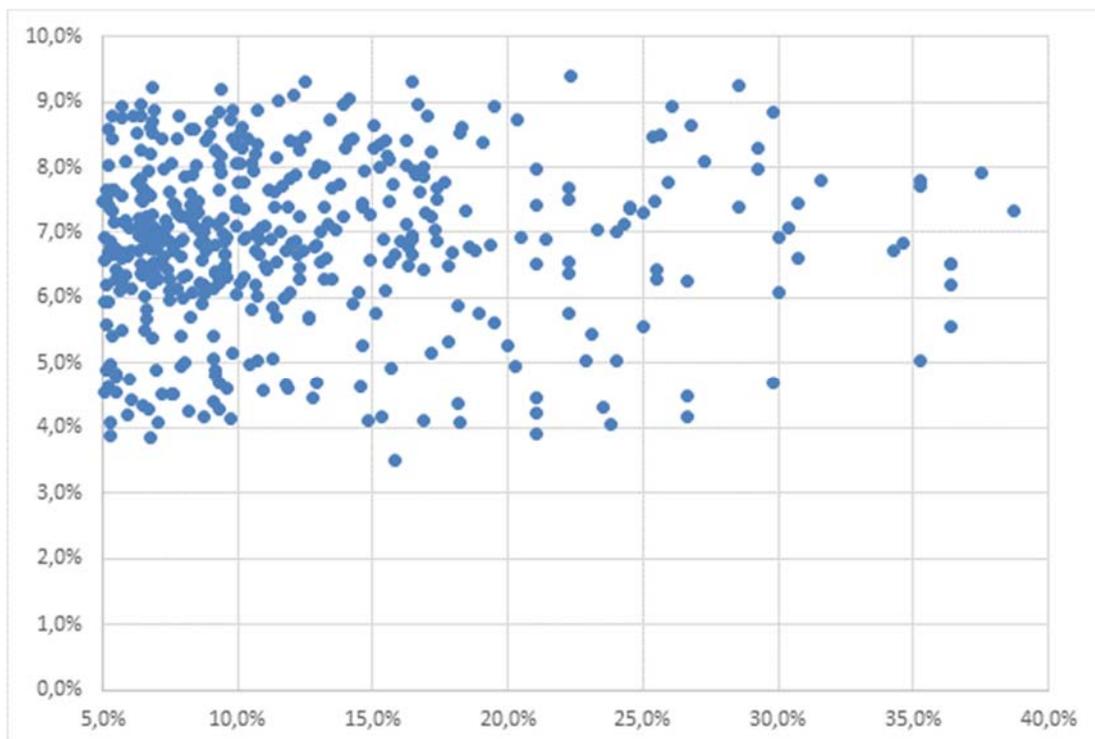
Bei den transformierten Werten lässt sich ein deutlicher Zusammenhang erkennen, der nach **Rück-Transformation** aber deutlich weniger markant hervortritt, siehe dazu auch die nachfolgende Abbildung.



**Abbildung 11: Beobachtung vs. Modell – gewichtetes additives Modell (2).**

Die Häufung der Beobachtungswerte bei null liegt daran, dass sehr häufig kein Storno beobachtet wurde, beispielsweise bei sehr wenigen Jahreseinheiten. Bei einem gewichteten Modell ist bei diesen Datenpunkten die Modellanpassung relativ beliebig, da sie ja gerade nicht ins Gewicht fällt. Die Beobachtungswerte bei 100% ergaben sich durch die Begrenzung der fiktiv erzeugten Werte, die teilweise über 100% lagen, wenn sich z. B. ein Storno auf 0,4 Jahreseinheiten bezieht.

Da die meisten Beobachtungswerte unter 40% liegen und die Kumulierung bei 0% keine weitere Aussagekraft hat als dass der Bestand zu gering war, ist nachfolgend noch einmal für die **rück-transformierten** Werte der Plot im Bereich von 5% bis 40% dargestellt.



**Abbildung 12: Beobachtung vs. Modell – gewichtetes additives Modell (3).**

Man kann einen leichten Trend im unteren Bereich bis ca. 15% erkennen, der aber nicht sehr spezifisch ist.

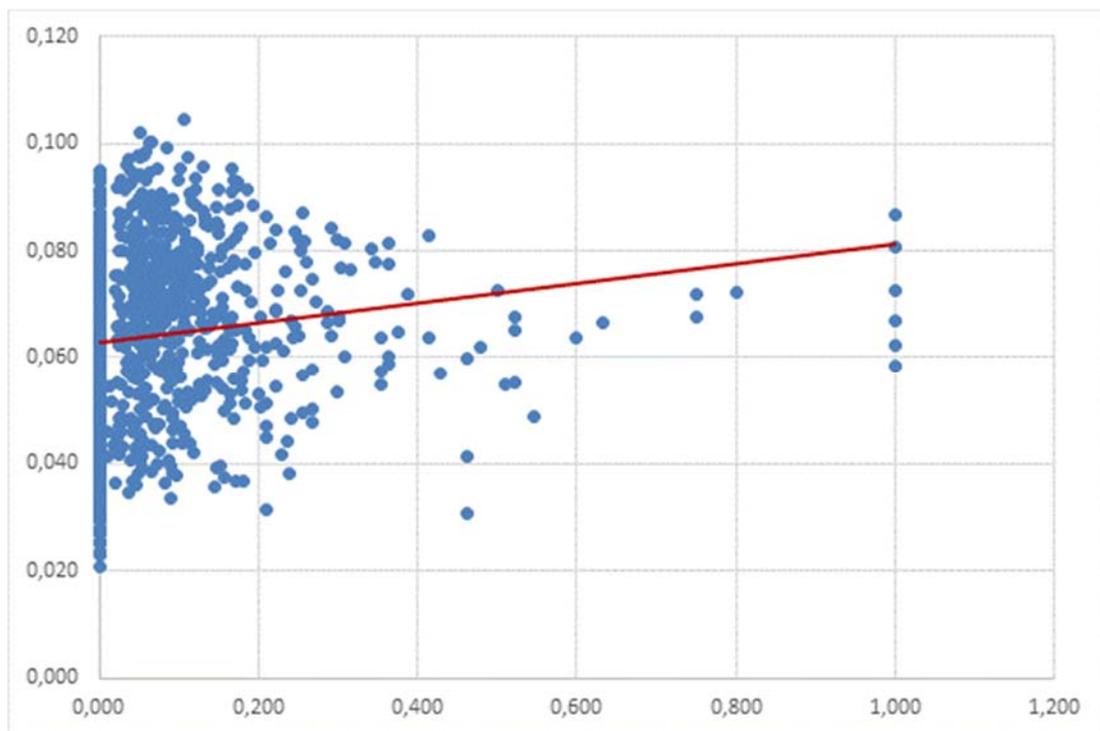
## 4.2 Ungewichtetes additives Modell

Die wichtigsten Eckdaten des ungewichteten additiven Modells sind nachfolgend zusammengefasst:

Fehlererwartungswert	0,000
Fehlervarianz	0,013
Abweichung Erwartungstreue in %	0,0%
Minimum modellierte Werte	2,1%
Maximum modellierte Werte	10,4%
Fehlervarianz modellierte Werte	0,013
Korrelation transformierte Werte	13,6%

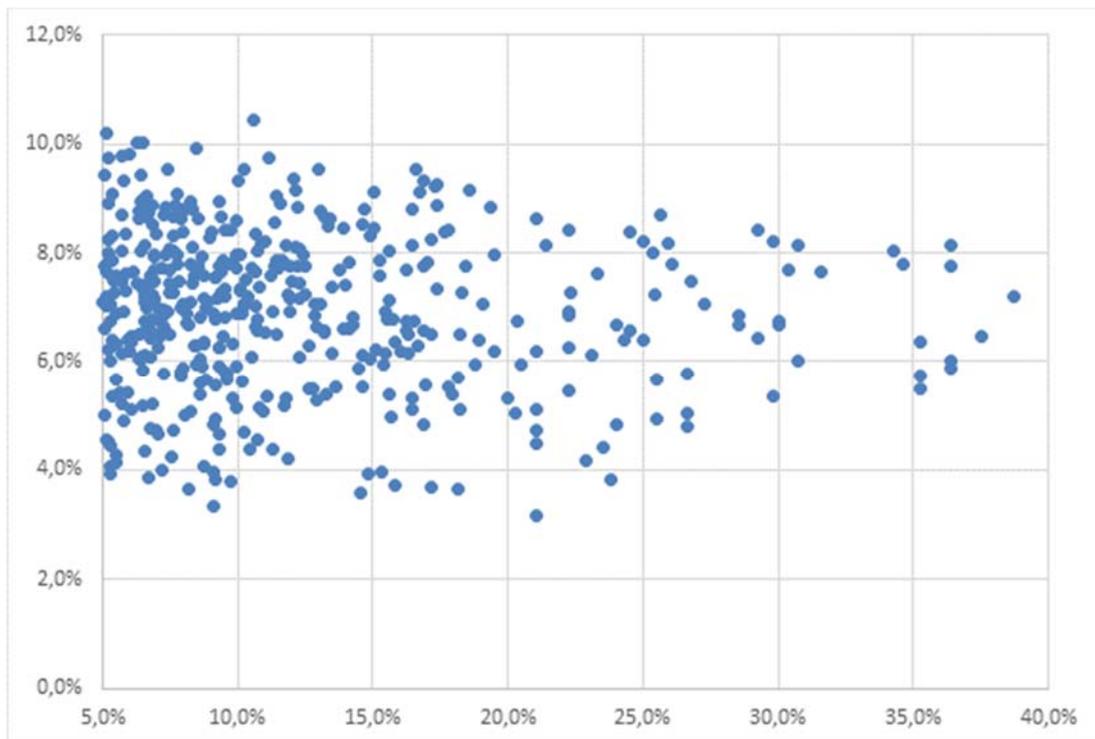
Die Fehlervarianz fehlt deutlich höher aus als beim gewichteten Modell, wobei die modellierten Werte etwas deutlicher spreizen, aber immer noch in einer überschaubaren Bandbreite liegen. Da hier keine Transformation durchgeführt wurde, unterscheiden sich die verschiedenen Korrelationen nicht, so dass hier nur ein Wert angegeben wurde. Die Korrelation ist deutlich kleiner als zuvor.

Die beobachteten versus modellierten Werte sind in der nachfolgenden Abbildung geplottet, wobei die Regressionsgerade in die geplotteten Daten eingefügt wurde. Die Häufung bei null ergibt sich wieder aus den vorinitialisierten Werten.



**Abbildung 13: Beobachtung vs. Modell – ungewichtetes additives Modell (1).**

Für eine bessere Übersicht ist wieder auch der Bereich von 5% bis 40% dargestellt ist, siehe dazu die nachfolgende Abbildung.



**Abbildung 14: Beobachtung vs. Modell – ungewichtetes additives Modell (2).**

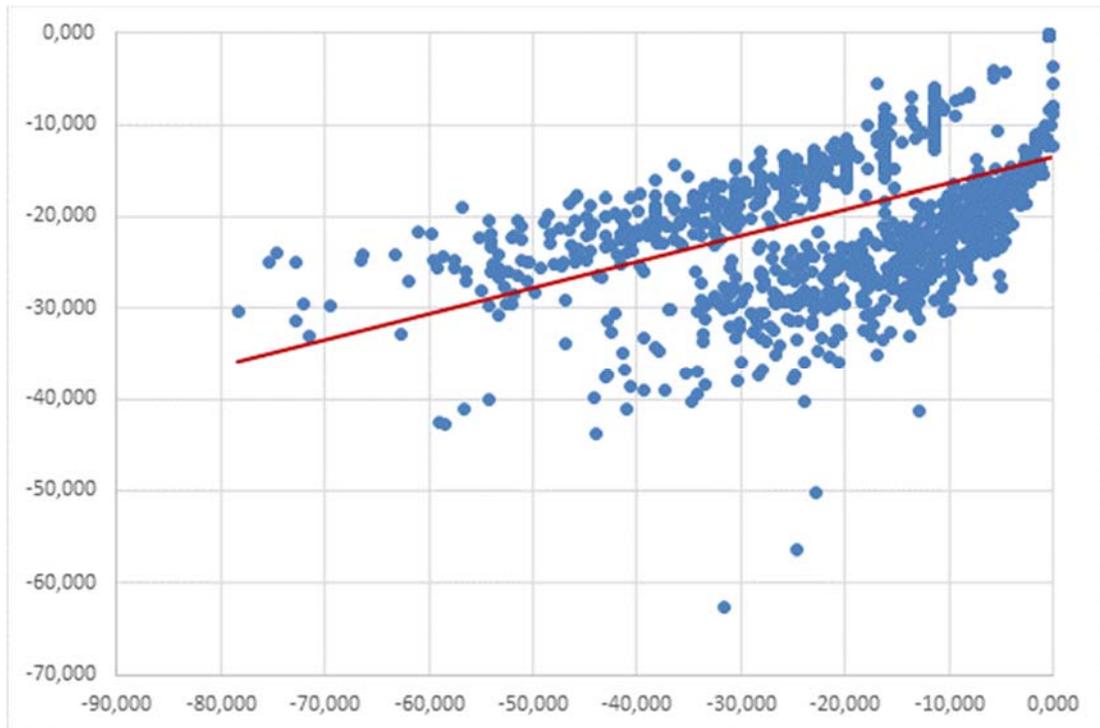
Der Zusammenhang ist noch unspezifischer und weniger bis gar nicht erkennbar als beim gewichteten Modell.

### 4.3 Gewichtetes multiplikatives Modell

Zunächst einmal sind auch für dieses Modell die wichtigsten Kenngrößen nachfolgend zusammengefasst:

Gewichteter Fehlererwartungswert	0,000
Gewichtete Fehlervarianz	6,208
Abweichung Erwartungstreue in %	27,7%
Minimum modellierte Werte	0,0%
Maximum modellierte Werte	27,8%
Fehlervarianz modellierte Werte	0,006
Korrelation transformierte Werte	48,4%
Korrelation Ausgangswerte	3,3%
Korrelation Ausgangswerte gewichtet	12,0%

Die gewichtete Fehlervarianz für die Modellvariable ist jetzt deutlich höher als bei den additiven Modellen, da nach dem Logarithmieren jetzt auch sehr stark negative Werte auftreten, insbesondere durch die Gewichtung. Aufgrund dieser hohen Varianz bei den (gewichteten) Logarithmen weicht das Modell sehr stark von der Erwartungstreue ab, so dass man entsprechend modifizieren muss. Nach der Modifikation ist die Fehlervarianz bei den Stornowahrscheinlichkeiten allerdings wieder sehr gering.

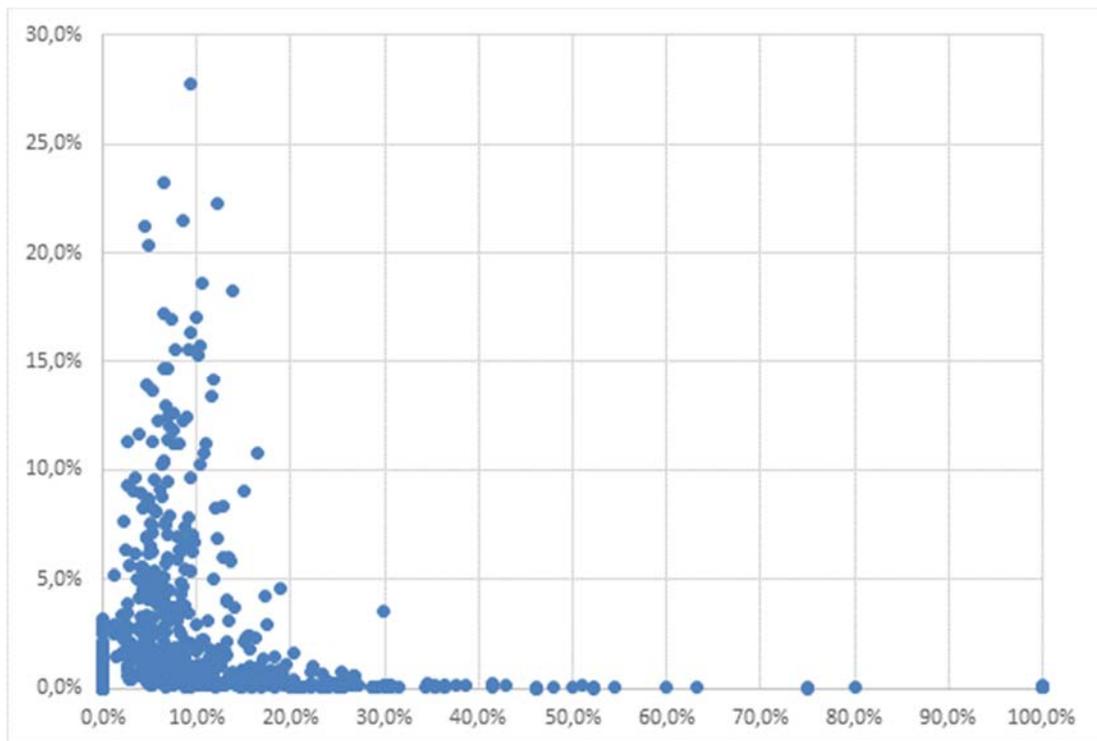


**Abbildung 15: Beobachtung vs. Modell – gewichtetes multiplikatives Modell (1).**

Bei den **transformierten Werten** ergibt sich eine erhöhte (wenn auch nicht besonders hohe) Korrelation, wobei der Zusammenhang in den geplotteten Daten durchaus erkennbar ist, siehe auch die obige Abbildung.

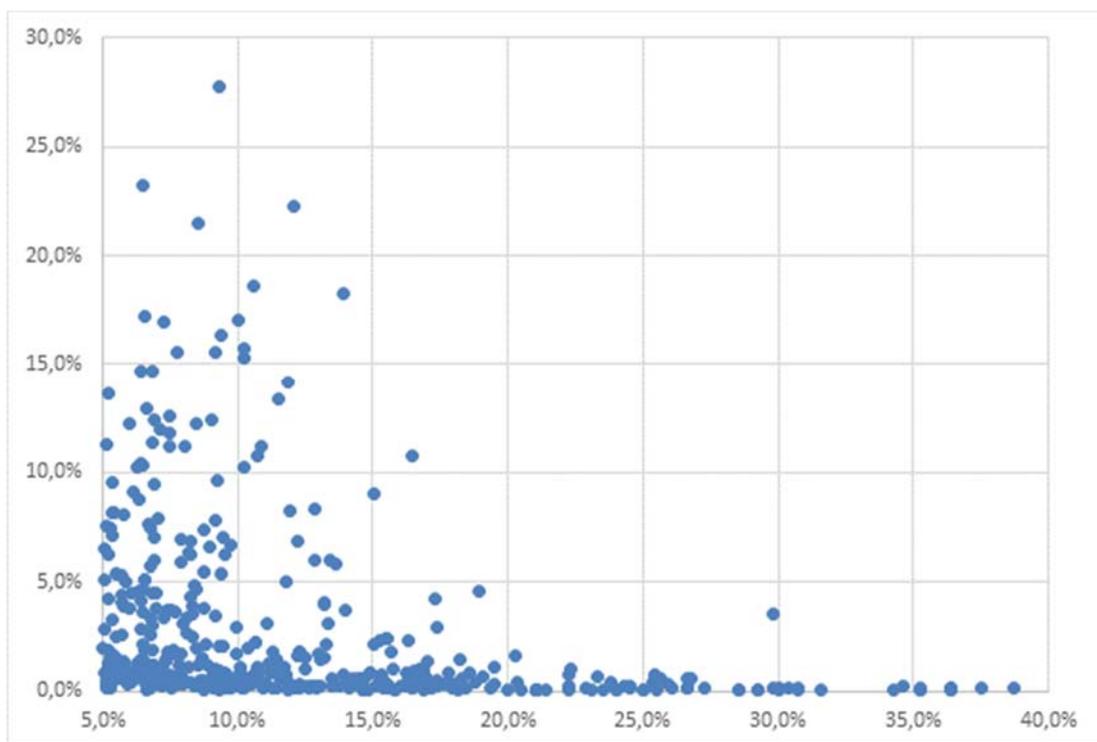
Die modifizierten modellierten Werte spreizen deutlich stärker als bei den additiven Modellen, wobei jetzt der Wert null als Minimum auch ziemlich nah erreicht wird. Die gewichtete Korrelation ist ähnlich gering wie beim ungewichteten additiven Modell und deutlich geringer als bei den transformierten Werten, für die die Ausgleichung ja durchgeführt wurde.

Die geplotteten Werte für die **beobachteten Ausgangswerte** mit den rücktransformierten (und im Hinblick auf Erwartungstreue angepassten) Ausgleichswerten sind in der nachfolgenden Abbildung dargestellt.



**Abbildung 16: Beobachtung vs. Modell – gewichtetes multiplikatives Modell (2).**

Wie man erkennen kann, ist die Modellierung systematisch „verzerrt“. Trotz einer geringen positiven Korrelation gibt es viele kleine Beobachtungswerte (ungleich der initialisierten Werte), denen hohe modellierte Werte zugeordnet werden, während es auch viele hohe Beobachtungswerte gibt, denen modellierte Werte nahe bei null zugeordnet werden.



**Abbildung 17: Beobachtung vs. Modell – gewichtetes multiplikatives Modell (3).**

In der obigen Abbildung erkennt man gut, dass sich diese Verzerrung auch nicht „bessert“, wenn man den Bereich einschränkt. Es gibt „systematisch“ einen Bereich, der nach der Modellierung sehr nahe bei null liegt. Daraus resultiert wahrscheinlich auch die starke Unterschätzung der beobachteten Stornowahrscheinlichkeiten.

#### 4.4 Ungewichtetes multiplikatives Modell

Wie man aus der nachfolgenden Übersicht der wichtigsten Eckwerte erkennt, ist dieses Modell in jeder Sicht „extrem“.

Gewichteter Fehlererwartungswert	0,000
Gewichtete Fehlervarianz	14,003
Abweichung Erwartungstreue in %	76,9%
Minimum modellierte Werte	0,0%
Maximum modellierte Werte	105,0%
Fehlervarianz modellierte Werte	0,021
Korrelation transformierte Werte	58,1%
Korrelation Ausgangswerte	2,5%
Korrelation Ausgangswerte gewichtet	2,5%

Die Fehlervarianz der Modellvariablen ist so hoch wie in keinem der anderen Modelle, woraus eine extreme Abweichung bei der Erwartungstreue resultiert. Nach der Modifikation ergeben sich daher modifizierte modellierte Stornowahrscheinlichkeiten, die über 100% liegen. Das ist zwar theoretisch bei zu geringer Bestandsmasse beobachtbar, sollte aber in einem guten Modell „weg geglättet“ werden.

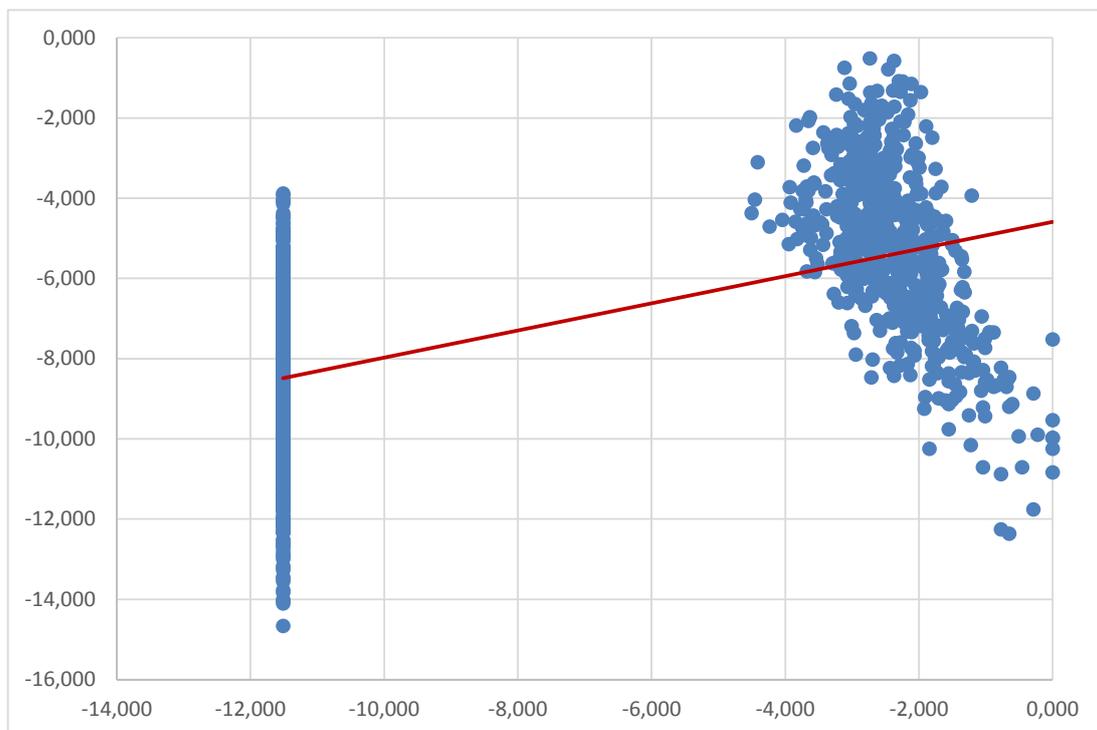
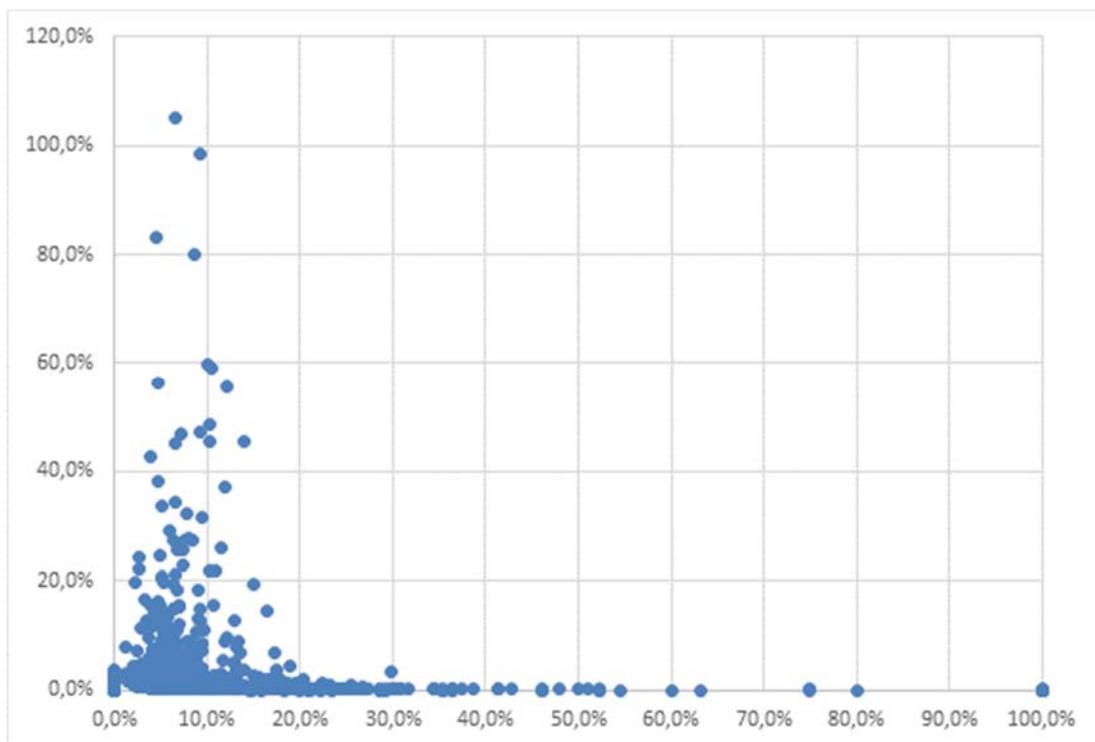


Abbildung 18: Beobachtung vs. Modell – ungewichtetes multiplikatives Modell (1).

Zwar gibt es eine deutliche (wenn auch nicht übertrieben hohe) Korrelation bei den **transformierten Werten** (siehe dazu die obige Abbildung), der Plot wirkt aber „irritierend“. Ohne die deutlich erkennbaren initialisierten Ausgangswerte würde man rein „optisch“ fast schon einen negativen Zusammenhang erwarten.

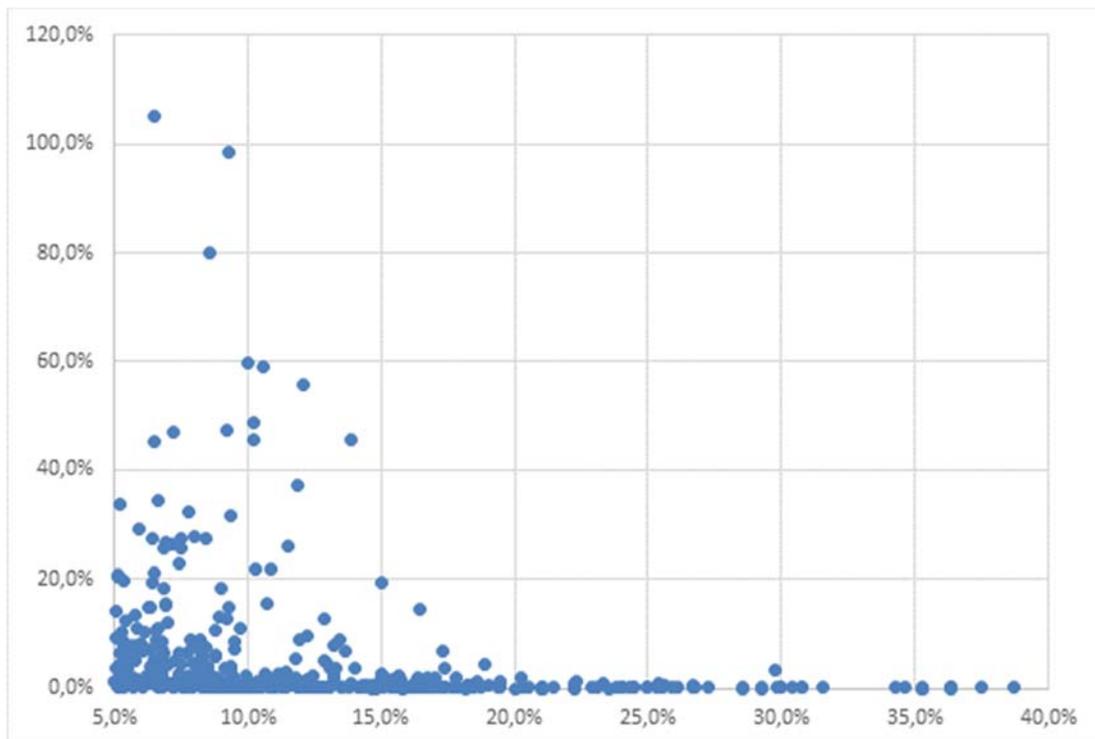
Der „Zusammenhang“ ergibt sich hier aus der Tatsache, dass zwei völlig verschiedene Cluster (die initialisierten Werte als Cluster 1 ohne jeden Zusammenhang und alle anderen Werte als Cluster 2 mit eher negativem Zusammenhang) auftreten, deren mittlere Niveaus ansteigen. Aufgrund der fehlenden Gewichtung werden alle Werte gleich behandelt und der „Niveaustieg“ als positiver Zusammenhang interpretiert.

Wie man aus der nachfolgenden Abbildung erkennt, ist eine Korrelation zwischen modellierten (und im Hinblick auf eine Erwartungstreuung angepassten) und beobachteten **Ausgangswerten** bei diesem Modell nicht mehr erkennbar, d. h. durch das Modell ergibt sich eigentlich keine Verbesserung beispielsweise im Vergleich zu einer frei ausgewürfelten modellierten Stornowahrscheinlichkeit.



**Abbildung 19: Beobachtung vs. Modell – ungewichtetes multiplikatives Modell (2).**

Die extremen Eigenschaften „verbessern“ sich auch nicht, wenn man den Beobachtungsbereich des Plots einschränkt.



**Abbildung 20: Beobachtung vs. Modell – ungewichtetes multiplikatives Modell (3).**

Das ungewichtete multiplikative Modell ist also in jeder Hinsicht ungeeignet zur Ausglei-  
 chung von beobachteten Stornowahrscheinlichkeiten.

Zusammenfassend kann man sagen, dass nur das gewichtete additiven Modell als geeignet  
 erachtet werden kann, dies umso mehr, da es keine negativen modellierten Stornowahr-  
 scheinlichkeiten gibt. Der Korrelationszusammenhang ist nach Rücktransformation aller-  
 dings nicht sehr hoch.

## 5 Fazit

Gauß Markow Modelle stellen eine Möglichkeit dar, den bekannten Ansatz der linearen Regression als Ausgleichsverfahren mit Hilfe einer Geraden durch die Datenpunkte zu erweitern.

Im Unterschied zu verallgemeinerten linearen Modellen sind Gauß Markow Modelle noch vergleichsweise einfach umzusetzen, beispielsweise in EXCEL mit Hilfe von Visual Basic. Numerische Probleme können mit Hilfe der Cholesky Zerlegung i. d. R. ganz gut vermieden werden.

Die einfache Umsetzung erfordert aber ihren Preis. So steckt in dem Modell implizit die Annahme eines normalverteilten Fehlers, da unter dieser Annahme die Methode der kleinsten Quadrate einen Maximum Likelihood Schätzer liefert. Bei anderen Fehlerverteilungen ist das nicht unbedingt der Fall.

Sofern es einen Offsetvektor in der Designmatrix gibt, erhält man die minimale Fehlervarianz, da der Erwartungswert der Fehler null beträgt. (Das Modell ist also in diesem Fall erwartungstreu.) Sofern zweckmäßig kann man dabei durchaus auch mit einer Gewichtung arbeiten; man erhält dann die minimale gewichtete Fehlervarianz bzw. einen gewichteten Erwartungswert von null für den Fehler.

Gauß Markow Modelle sind von der Konzeption her additive Modellansätze. Durch Transformieren der Beobachtungen kann man zwar multiplikative Modelle konzipieren; diese sind allerdings dann nicht mehr notwendigerweise erwartungstreu. Wenn das Modell vergleichsweise schlecht zu den Daten passt, sind die Abweichungen u. U. erheblich.

Mit Gauß Markow Modellen kann man die lineare Regression auf Ausgleichungen höherer Ordnung erweitern; die Lösungen ergeben sich hier direkt aus dem allgemeinen Lösungsansatz. Interpretationen als Momente der Verteilung benötigt man hier nicht mehr.

Im Unterschied zu verallgemeinerten linearen Modellen liefern Gauß Markow Modelle nicht notwendigerweise eine Signifikanzaussage. Das bestmögliche Modell muss nicht unbedingt ein gutes Modell sein. Gewählte Modellierungsparameter werden nicht im Hinblick auf ihre Modellaussagekraft getestet.

Im Prinzip war dieses Problem bereits bei der linearen Regression aufgetaucht. Bei geringer Korrelation hat die ausgeglichene Gerade wenig Aussagekraft. Aber auch eine hohe Korrelation muss nicht auf eine gute Ausgleichung hinweisen. So wird bei einem exponentiellen Zusammenhang die Korrelation i. d. R. deutlich positiv sein, da sowohl die Gerade als auch die Exponentialfunktion mit sicherer Tendenz ansteigen. Innerhalb des Modellierungsbereichs (Interpolation) ist die Ausgleichung daher i. d. R. ganz gut; außerhalb des Modellierungsbereichs (Extrapolation) verliert die Ausgleichung aber systematisch an Qualität, da die Exponentialfunktion jeder Gerade „davonrennt“.

Gauß Markow Modelle darf man daher nicht einfach automatisch anwenden, sondern man muss immer auch testen (ggf. durch einfache Visualisierungstechniken), ob das gewählte Modell auch adäquat zum Problem ist.

## Literaturverzeichnis

- [1] Koch: Parameterschätzungen und Hypothesentests in linearen Modellen. Dümmler, Bonn, 1980.
- [2] Scuzzarello: Stornoanalyse bei einem Schadenversicherungsbestand. Master Arbeit, Institut für Versicherungswesen, Fakultät für Wirtschafts- und Rechtswissenschaften, Fachhochschule Köln, Köln, 2014.
- [3] Kahlenberg: Storno und Profitabilität in der Privathaftpflichtversicherung. Eine Analyse unter Verwendung von univariaten und bivariaten verallgemeinerten linearen Modellen. Aachen, Shaker Verlag, 2005.

## Abbildungsverzeichnis

Abbildung 1: Auszug aus der komprimierten Beobachtungsdatei. ....	10
Abbildung 2: Auszug aus den transformierten Werten – gewichtetes additives Modell. ....	11
Abbildung 3: Auszug aus der transformierten Designmatrix – gewichtetes Modell.....	12
Abbildung 4: Transformierte Lösungsmatrix – gewichtetes additives Modell. ....	12
Abbildung 5: Cholesky Zerlegung – gewichtetes additives Modell (1).....	13
Abbildung 6: Cholesky Zerlegung – gewichtetes additives Modell (2).....	13
Abbildung 7: Auszug aus den modellierten Werten – gewichtetes additives Modell. ....	14
Abbildung 8: Auszug aus den modellierten Werten – gewichtetes multiplikatives Modell.	15
Abbildung 9: Auszug aus den angepassten Werten – gewichtetes multiplikatives Modell.	16
Abbildung 10: Beobachtung vs. Modell – gewichtetes additives Modell (1).....	18
Abbildung 11: Beobachtung vs. Modell – gewichtetes additives Modell (2).....	18
Abbildung 12: Beobachtung vs. Modell – gewichtetes additives Modell (3).....	19
Abbildung 13: Beobachtung vs. Modell – ungewichtetes additives Modell (1). ....	20
Abbildung 14: Beobachtung vs. Modell – ungewichtetes additives Modell (2). ....	21
Abbildung 15: Beobachtung vs. Modell – gewichtetes multiplikatives Modell (1). ....	22
Abbildung 16: Beobachtung vs. Modell – gewichtetes multiplikatives Modell (2). ....	23
Abbildung 17: Beobachtung vs. Modell – gewichtetes multiplikatives Modell (3). ....	23
Abbildung 18: Beobachtung vs. Modell – ungewichtetes multiplikatives Modell (1).....	24
Abbildung 19: Beobachtung vs. Modell – ungewichtetes multiplikatives Modell (2).....	25
Abbildung 20: Beobachtung vs. Modell – ungewichtetes multiplikatives Modell (3).....	26

## Impressum

Diese Veröffentlichung erscheint im Rahmen der Online-Publikationsreihe „Forschung am IVW Köln“. Alle Veröffentlichungen dieser Reihe können unter [www.ivw-koeln.de](http://www.ivw-koeln.de) oder [hier](#) abgerufen werden.

**Forschung am IVW Köln, 8/2014**

**Heep-Altiner, Münchow, Scuzzarello: Ausgleichsrechnungen mit Gauß Markow Modellen am Beispiel eines fiktiven Stornobestandes**

**Köln, September 2014**

**ISSN (online) 2192-8479**

Herausgeber der Schriftenreihe / Series Editorship:

Prof. Dr. Lutz Reimers-Rawcliffe  
Prof. Dr. Peter Schimikowski  
Prof. Dr. Jürgen Strobel

Institut für Versicherungswesen /  
Institute for Insurance Studies

Fakultät für Wirtschafts- und Rechtswissenschaften /  
Faculty of Business, Economics and Law

Fachhochschule Köln / Cologne University of Applied Sciences

Web [www.ivw-koeln.de](http://www.ivw-koeln.de)

Schriftleitung / Contact editor's office:

Prof. Dr. Jürgen Strobel

Tel. +49 221 8275-3270  
Fax +49 221 8275-3277

Mail [juergen.strobel@fh-koeln.de](mailto:juergen.strobel@fh-koeln.de)

Institut für Versicherungswesen /  
Institute for Insurance Studies

Fakultät für Wirtschafts- und Rechtswissenschaften /  
Faculty of Business, Economics and Law

Fachhochschule Köln / Cologne University of Applied Sciences  
Gustav Heinemann-Ufer 54  
50968 Köln

Kontakt Autor / Contact author:

Prof. Dr. Maria Heep-Altiner  
Institut für Versicherungswesen /  
Institute for Insurance Studies

Fakultät für Wirtschafts- und Rechtswissenschaften /  
Faculty of Business, Economics and Law

Fachhochschule Köln / Cologne University of Applied Sciences  
Gustav Heinemann-Ufer 54  
50968 Köln

Tel. +49 221 8275-3449  
Fax +49 221 8275-3277

Mail [maria.heep-altiner@fh-koeln.de](mailto:maria.heep-altiner@fh-koeln.de)