Anne Thomas

Fuzzy Set Theory with Applications in Claims Reserving

Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag

Fuzzy Set Theory with Applications in Claims Reserving

Anne Thomas

Fuzzy Set Theory with Applications in Claims Reserving

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

1. Aufl. - Göttingen: Cuvillier, 2017
Zugl.: Hamburg, Univ., Diss., 2016

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2017
Gedruckt auf umweltfreundlichem, säurefreiem Papier aus nachhaltiger Forstwirtschaft.

ISBN 978-3-7369-9597-0
eISBN 978-3-7369-8597-1

Insofern sich die Sätze der Mathematik auf die Wirklichkeit beziehen, sind sie nicht sicher, und insofern sie sicher sind, beziehen sie sich nicht auf die Wirklichkeit.

Albert Einstein[1]

[1] Albert Einstein, "Geometrie und Erfahrung", Festvortrag am 27. Januar 1921 in der preußischen Akademie der Wissenschaften in Berlin

Acknowledgements

I am indebted to my supervisor Prof. Dr. Michael Merz who drew my attention to the applications of fuzzy methods in the field of claims reserving and helped to start exploring this “island”. I want to thank him for exchanging ideas, critical explanatory notes, and motivating talks. He also gave me the opportunity to present the current findings at international conferences, among them the ASTIN Colloquium 2013, the IFSA World Congress/NAFIPS Annual Meeting 2013, the APRIA Conference 2013 and the 18th International Congress on Insurance: Mathematics & Economics 2014. I would also like to thank the participants for valuable feedback and advice.

I am grateful to Prof. Dr. Bernhard Arnold for being my second supervisor and for his help in my early steps in the “fuzzy world”. I thank Prof. Dr. Wolfgang Drobetz for chairing the examination committee.

Thanks go to my colleagues at the chair of Mathematics & Statistics in Business Administration Nha-Nghi de la Cruz, Tobias Gummersbach, Sebastian Happ, Jochen Heberle, Marie Hielscher and Arne Johannssen. Particular notice deserves Jochen Heberle with whom I enjoyed productive collaborations and Sebastian Happ for proofreading drafts of this work as well as valuable discussions. Moreover, I would like to thank my colleagues of the Institute of Econometrics and Statistics for a pleasant time.

I would like to thank Angelika Ruiz Merino for her constant commitment to the chair of Mathematics & Statistics in Business Administration and enjoyable lunch breaks.

I thank my family for their continuous belief in me, their encouragements and distractions from the academic world. Last but not least I like to thank Ole for being a resting pole in times of ups and downs.

Anne Thomas

Contents

List of Figures

List of Tables

List of Symbols

$\oplus$	Fuzzy addition
$\ominus$	Fuzzy subtraction
$\otimes$	Fuzzy multiplication
$\oslash$	Fuzzy division
$\mathbf{0}_{k\times l}$	Zero matrix with k rows and l columns
$\tilde{A}$	Fuzzy set
$\tilde{A}_\alpha$	α-level set of a fuzzy set $\tilde{A}$
$\tilde{a}$	Fuzzy number
$(a, l_a, r_a)_{L,R}$	Fuzzy number of LR-type
$\boldsymbol{\alpha}^*$	α-pattern, i.e. vector containing H-values for every single fuzzy regression
$C_{i,j}$	Cumulative claims made in relative accident year i and relative development year j
$\tilde{C}_{i,j}$	Cumulative claim considered as a fuzzy number
COG	Center of gravity
$\mathcal{D}_I$	Set of observations
$E_\beta(\tilde{a})$	Expected value of a fuzzy number $\tilde{a}$
$F_{i,j}$	Individual claims development factor
f_j	Development factors
$\tilde{f}_j^{\text{AFCL}} = \left(f_j^{\text{AFCL}}, l_{f_j^{\text{AFCL}}}, r_{f_j^{\text{AFCL}}}\right)$	Fuzzy chain-ladder factor in model with fuzzy regression approach (Chapter 6)
$F\mathscr{P}(X)$	Fuzzy power set
$(\gamma_j)_{j\in\{0,\ldots,J\}}$	Claims development pattern
$\tilde{\gamma}_j = (\gamma_j, l_{\gamma_j}, r_{\gamma_j})$	Fuzzy claims development pattern in the fuzzy Bornhuetter Ferguson method
$\text{hgt}(\tilde{A})$	Height of a fuzzy set $\tilde{A}$
$\mathbf{I}_n$	Identity matrix with n rows and columns
L, R	Reference functions

λ_{i+j}	Factor describing calendar year effects in Taylor's geometric separation method
M	Measure of fuzziness
$\mu_{\tilde{A}}$	Membership function of a fuzzy set $\tilde{A}$
N_i	Number of claims occurred in accident year i
ν_i	A priori information in the Bornhuetter Ferguson method
$\tilde{\nu}_i = (\nu_i, l_{\nu_i}, r_{\nu_i})$	Fuzzy a priori information in the fuzzy Bornhuetter Ferguson method
$(\Omega, \mathcal{A}, P)$	Probability space
P	Probability measure
P_j	Share of claims settled in development year j in Taylor's geometric separation method
$\mathscr{P}(X)$	Power set of X
R	Aggregated reserve
R_i	Reserve for accident year i
$\tilde{R}_i$	Fuzzy reserve
$\mathrm{supp}(\tilde{A})$	Support of a fuzzy set $\tilde{A}$
$\mathrm{Unc}_K(\tilde{a})$	Uncertainty of a triangular fuzzy number $\tilde{a}$
$\mathbf{W}$	Weighting matrix
$X_{i,j}$	Incremental claim in relative accident year i and relative development year j
$Z_{i,j}$	Claims payments

List of Abbreviations

ANOVA	Analysis of variance
BF	Bornhuetter Ferguson
BLU	Best linear unbiased
CAPM	Capital asset pricing model
CL	Chain-ladder
CWW	Computing with words
e.g.	for example
FBF	Fuzzy Bornhuetter Ferguson
FCL	Fuzzy chain-ladder
FN	Fuzzy number
FR	Fuzzy regression
FST	Fuzzy set theory
GLM	Generalized linear model
IBNR	Incurred but not reported
IBNER	Incurred but not enough reserved
i.e.	that is
i.i.d.	independently and identically distributed
LoB	Line of business
LP	Linear program
LSRM	Linear stochastic reserving method
MSEP	Mean squared error of prediction
OLS	Ordinary least squares
PIC	Paid incurred chain
se	Standard error
s.t.	subject to
TFN	Triangular fuzzy number

1 Introduction

Inaccuracies are omnipresent in our every day life ranging from our daily language to sketchy decision criteria to errors of measurement. Already Oskar Morgenstern[2] pointed out that all economic decisions are characterized by the fact that both quantitative and non-quantitative information influence the act of decision-making (cf. Morgenstern 1965, p. 1). Even emotions and mainly intuition can be of central importance in the process (cf. Holtfort 2011, p. 507).

Intuition can be regarded in different ways, one of which is considering it as knowledge gained by experience. Players knowingly and unknowingly gather information and keep them in mind. That is why they make decisions "from the gut" instead of employing a process of weighing up different alternatives (cf. Holtfort 2011, p. 508). Also economic theory deals with contents which are e.g. obtained from personal experiences or other "non-constructional" ways (cf. Morgenstern 1965, p. 88). Hence, subjective judgment can be an important factor.

Another matter that needs to be raised concerning impreciseness are linguistic inaccuracies which are present in our daily language. People speak of something being "likely" or "unlikely", that a "large" claim occurred, that an event will happen in the "near" future, etc. Usually, the terms are not precisely mathematically described and implemented in the decision-making process or mathematical models but remain vague.

An actuary at an insurance company is faced with similar situations. On the one hand, his/her calculations and the underlying models need to be as precise as possible and should correspond to the given data whether it is in the field of pricing, claims reserving, product controlling or other. At the same time, he/she should form an opinion or assessment to what extent his/her calculations express a realistic view of (future) reality.

[2]Oskar Morgenstern, * 1902, † 1977, was an Austrian-US-American economist.

Adjustments due to personal opinions are necessary or common in various situations: Circumstances are thinkable in which the given data is not sufficient or does not reflect a realistic view. Whether the derived consequences are too optimistic or pessimistic – in comparison to personal experience – adjustments in one or the other direction are possible.
Moreover, personal assessment can be found in the context of risk classification. "Determining the significant factors [...] however, may involve some subjective judgment" (cf. Outreville 1988, p. 150). Risks are often clustered in risk classes. If a risk can be seen as part of one class or another, an actuary will decide to which class the risk belongs. The presence of subjective judgment can also be seen in connection to pricing within the use of expert systems e.g. when historical data and/or information is not available (cf. Culp 2006, p. 164). In claims reserving subjective judgment is also present when estimating loss reserves on a single case basis (cf. Lemaire 1988, p. 396). In this particular case, actuaries often do not employ mathematical models but set up reserves according to the information they have at hand as well as their personal opinion and experience.

In practice, the actuary tends to adjust particular data (or parts of it) solely based on his/her intuition. This also applies in the field of claims reserving on which the present work will focus. This subdiscipline of insurance mathematics provides (stochastic) methods which are methodically advanced and diverse. The literature on claims reserving deals with various models which differ in a lot of facets: whether they are purely computational or make use of a stochastic framework, employ credibility theory or the Kalman filter, refer to paid or incurred data, etc. In addition, regulatory frameworks set more legal requirements in order to determine the reserve.

Nevertheless, the methods presented in the literature do not generally comprise a formalized approach to consider these subjective judgments. From a methodological point of view, the theory of fuzzy set seems to be appropriate. Fuzzy sets represent an approach to describe vagueness and imprecision. Particularly, it aims to depict limitations on knowledge and inaccuracies in context with numerical descriptions of problems. Thereby, it emanates from subjective judgments and not from global assumptions. Since its introduction in the seminal paper of Lotfi A. Zadeh (cf. Zadeh 1965) it underwent a continuous further development. It provides means to compute with vague expressions and a whole "computing tool box" has been developed. It took about 20 years until the first applications in insurance were discussed and another 20 years till the theory found its way into claims reserving.

In view of the above, the present work pursues two main goals. On the one hand, it aims to give an overview of existing methodical approaches for depicting subjective judgment in claims reserving. In this context, all necessary basics of the two branches of research, i.e. claims reserving and fuzzy set theory, are summarized. On the other hand, the present work will show in what way subjective judgment can be implemented in the two most popular claims reserving methods, i.e. the chain-ladder (CL) (cf. Section 4.3) and the Bornhuetter Ferguson (BF) (cf. Section 4.4) method. Therefore, three new approaches – two based on CL and one based on BF – are developed in the scope of this dissertation.

As already mentioned, subjective judgment can arise in the field of claims reserving. This is addressed in Chapter 5 of which an earlier version has been published in *Insurance: Mathematics & Economics* (cf. Heberle and Thomas 2016). The fundamental idea of the fuzzy chain-ladder (FCL) model is to model the observation that actuaries might adjust previously calculated development factors according to their subjective judgment in the context of the CL method using fuzzy numbers and their corresponding arithmetic. To the best of our knowledge, the CL method has not been the center of studies within fuzzy applications in spite of its popularity. The method aims to model the development factors with the help of triangular shaped fuzzy numbers with equal spreads as they are easy to implement. Moreover, the derived reserve has been fuzzy as well in earlier models, i.e. not a specific value has been derived but a range of possible values. Since a reserve is a figure in the balance sheet of an insurance company it needs to be specified as a crisp number. Therefore, our method will apply a defuzzification method and also will make an attempt to quantify the uncertainty of the prediction.

This idea shall be further pursued in the second approach (see Chapter 6) by applying a fuzzy regression method which assumes a functional relation described with fuzzy coefficients to the CL method in its linear model representation. In fact, a regression "tube" will be derived in which all data points lie. Methods of fuzzy regression partially provide possibilities to subjectively assess the scattering of the data around the regression line which describes a functional relationship. This fact shall be additionally taken into consideration, such that actuaries can subjectively assess the informative value of the data. For this approach we will model the fuzzy coefficients with triangular shaped fuzzy numbers of different spreads. The aim is to reduce the prediction uncertainty – depending on the data – in this model compared to the FCL model.

An additional idea is pursued in the third approach (see Chapter 7): The considered BF method utilizes a priori information which is given either by an external source (market statistics, expert knowledge, organizational data, etc.) or internally. In case a priori information is available the BF method is a popular alternative to the CL method. Depending on its origin the information can be afflicted with vagueness. The fuzzy Bornhuetter Ferguson (FBF) model aims to map this with the help of fuzzy numbers. Development factors as well as the a priori information shall be modeled with fuzzy numbers. Again, the goal is to derive predictions for the reserves.

The structure of the work is as follows: Subsequent to the general introduction the theory of fuzzy sets is introduced in Chapter 2. The basic concepts are presented and all necessary theory for the later chapters is provided. Moreover, a distinction between the theory of fuzzy sets and probability theory is drawn. The comparison is followed by a literature survey of applications of fuzzy methods in insurance in Chapter 3. It is classified by the field of application: underwriting, risk classification, pricing and claims reserving. The latter point is only mentioned briefly since it is discussed in detail in Chapter 4. Here the problem of claims reserving as well as an overview of reserving methods is addressed. Subsequently, the FCL method which makes use of fuzzy numbers and their corresponding arithmetic is presented in Chapter 5. The model is motivated and examined followed by a numerical example. As a further utilization a model of fuzzy regression is applied to the CL method in Chapter 6, in particular the representation as a sequence of linear models. The description of the methodology is followed by an example for which the same data base as before is used in order to draw a comparison. The FBF method is introduced in Chapter 7 and additionally takes into account a priori information. Within this model framework reserves and their corresponding uncertainty are deduced. A short summary is given and questions for further research are raised in Chapter 8.

2 | Fuzzy Theory

In reality we often face situations in which we cannot decide precisely whether an object belongs to a certain class or not. We might consider classes like “all tall men” or “all damages much higher than 10,000 €”. In these cases it is not always possible to distinguish between objects which are members of those classes and those which are not. A class like e.g. “all good students” is not a set in the mathematical sense as the word “good” is not described in a mathematically precise manner. A first approach to even consider those sets has been introduced by Zadeh (cf. Zadeh 1965).
The aim of this chapter is to give an introduction to the theory of fuzzy sets, fuzzy numbers and fuzzy regression. Firstly, we give a motivation for the introduction of this theory and outline the historic development in Section 2.1.1. Subsequently, a presentation of fuzzy sets according to Zadeh is given. In Section 2.2 fuzzy numbers as a special case of fuzzy sets are introduced and the corresponding arithmetic is defined. Finally, the basic concepts of fuzzy regression are presented in Section 2.3.

2.1 Fuzzy Sets

2.1.1 Historical Development

For a long time the scholars in philosophy taught that logic is a two-valued science. Therefore a proposition or statement is either e.g. “true” or “false”, “0” or “1”. Aristotle (384 BC – 322 BC), a student of Plato, was one of the founders of this school. Aristotle’s law of non-contradiction, the second law of “The Three Laws of Thought”, states that a proposition cannot be true and false at the same time. Additionally, his law of the excluded middle says that an object either possesses an attribute or its opposite.

Aristotle forbids that a proposition “in the middle” holds true, i.e. that an object only possesses an attribute to a certain extent (cf. Eberhart and Shi 2011, p. 284).

Scientists at the beginning of the twentieth century began to argue that not everything is two-valued. In mathematics a theory is developed for “perfect” objects. If objects fulfill these assumptions the proposition holds true, otherwise not. As objects in reality usually are not “perfect”, the mathematical findings are – strictly speaking – not applicable. Therefore, Black proposes to introduce a symbolism for vagueness, i.e. for the case if requirements are not completely fulfilled (cf. Black 1937, pp. 427-429).

In one of his lectures Bertrand Russell pursues a similar idea. “I propose that all language is vague and that therefore my language is vague” (cf. Russell 1923, p. 84). Language is often a source of problems in modeling. As an example the following situation could be considered: If a person tells his friend to be at his house shortly after five p.m., this can be perceived by his friend in another way than it was actually meant. In one understanding this might be the time interval from 5:00 to 5:10 p.m. whereas another person could be of the opinion that this is the time interval from 5:00 to 5:15 p.m. In order to cope with the problem of imprecise language Black proposed the use of so-called *consistency-profiles* to picture this vagueness (cf. Black 1937, pp. 430ff.).

A first idea of a set theory which allows for a multi-valued logic came up from Menger. His *ensembles flous* are the French counterpart to fuzzy sets (cf. Menger 1951). In fact, the work of Lotfi A. Zadeh laid the foundation of the further research on fuzzy sets (cf. Zadeh 1965) and his theory is closer to the idea of Black’s consistency-profiles (cf. Dubois and Prade 1980, p. 4).

After the publication of Zadeh’s pioneer article a fast development in research took place. A survey about the early works can be found in Kaufmann (1975). For further literature see e.g. Dubois and Prade (1980), Bandemer and Gottwald (1993), Zimmermann (2001) or Dubois and Prade (2000). Nonetheless, the reactions to Zadeh’s publication were divided. While the responses in the United States were quite cautious or there were harsh critics[3], scientists and technicians in Europe and Japan jumped onto the bandwagon. The use of fuzzy logic succeeded in Europe for the first time with the regulation of a steam raising unit in a power plant (cf. Altrock 1995, p. 7). Further applications in Japan were the regulation of a metro in which fuzzy technologies allowed for smooth

[3]For a collection of citations of critics see e.g. Altrock (1995, p. 7) and Dubois and Prade (2000, p. 3).

driveaways and slowdowns as well as implementations in camcorders, microwaves and washing machines (cf. Altrock 1995, pp. 8f.).

2.1.2 Basic Notations

The representation in this section basically relies on Bandemer and Gottwald (1993), Dubois and Prade (1980), Kruse et al. (1995) and Zimmermann (2001).
According to Cantor (1895) a set is "any collection M into a whole of definite, distinct objects m (which are called the "elements" of M) of our perception or of our thought" (cf. Cantor 1895, p. 481). In the following we will speak of a *crisp* set whenever we consider a set in the sense of Cantor. There are several ways to mathematically describe a set. Either every single element is listed or (for readability) the set is denoted with the help of descriptive characteristic traits, or the set is described with an indicator function (cf. Merz and Wüthrich 2013b, pp. 32ff.). As the objects are distinguishable in the case of crisp sets it is always possible to define an indicator function. Let Ω be a basic set and A a subset of Ω, i.e. $A \subset \Omega$. The function

$$\mathbf{1}_A : \ \Omega \longrightarrow \{0,1\}, \qquad \omega \mapsto \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

is called indicator function (or characteristic function) of A.

As stated before, there are linguistic inaccuracies in our everyday life as well as in business or scientific contexts. Indicator functions do not allow for gradual memberships to a certain set. In order to cope with that impreciseness Zadeh has defined the concept of fuzzy sets according to which we can specify the grade of membership for all elements $x \in X$ to a fuzzy set $\tilde{A}$ (cf. Zadeh 1965, p. 339). The idea was to enlarge the image set of a characteristic function and to permit grades of membership in the interval $[0,1]$.

Definition 2.1 (Fuzzy set)
Let $X \neq \emptyset$ be a collection of objects. A **fuzzy set** $\tilde{A}$ in X is defined by

$$\tilde{A} = \Big\{(x, \mu_{\tilde{A}}(x)) \mid x \in X\Big\}.$$

where $\mu_{\tilde{A}} : X \longrightarrow [0,1]$ is called **membership function**. We refer to $\mu_{\tilde{A}}(x)$ as the **grade of membership** of an element x in X with respect to $\tilde{A}$. Furthermore, we consider $F\mathscr{P}(X)$ as the **fuzzy power set** of X, i.e. the set of all fuzzy sets in X.

The membership function μ specifies to which extent an element belongs to a fuzzy set $\tilde{A}$. Consequently, the closer the grade of membership to one, the higher the grade of membership of x in $\tilde{A}$. Usually, those elements are assigned grade of membership of one which definitely belong to a set. The membership function is not necessarily bounded by zero and one. Nevertheless, it is a preferable setting. In fact, a fuzzy set is defined by its membership function.

An example of a membership function is shown in Figure 2.1. Dubois and Prade (1980) even refer to this kind of representation as extended Venn diagram. A classical Venn diagram cannot be drawn for fuzzy sets but an extended form can be used to visually verify set theoretic operations (cf. Dubois and Prade 1980, p. 14).

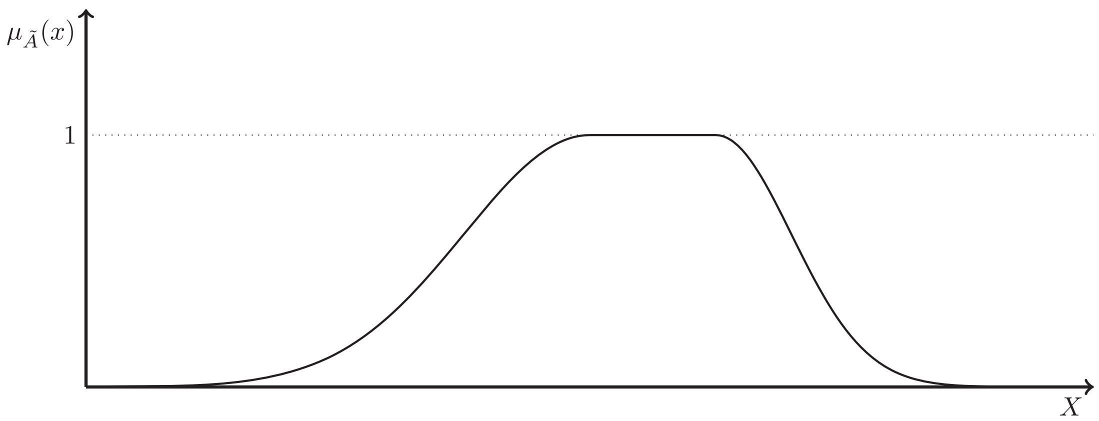


Figure 2.1: An example of a membership function of a fuzzy set $\tilde{A}$.

One criticism to fuzzy set theory often is the assignment of the grades of membership. In fact, they do not exist but are assigned to elements by an individual or a group. Therefore, it is a subjective assignment. There even might be fields of interest in which the mapping might be more controversial than in others. E.g. there might be more consensus in the field of “age” than in “smartness”.

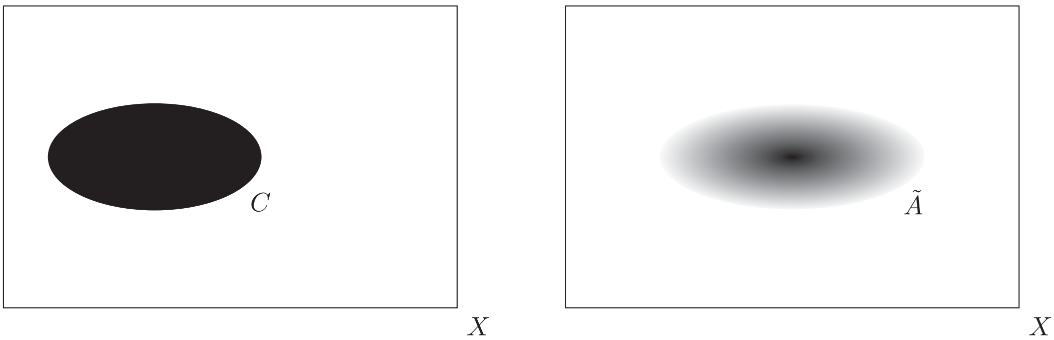
Remarks 2.2

a) The use of the interval $[0, 1]$ as the image set of the membership function allows for a convenient interpretation of the grade of membership.

b) Definition 2.1 ensures that every crisp set $A \in \mathscr{P}(X)$, where $\mathscr{P}(X)$ denotes the power set of X, can be considered as a fuzzy set since $\{0,1\} \subseteq [0,1]$.

c) Even though the membership function might be seen as a fuzzy analogon to a density function in probability theory there are remarkable differences which support that those two concepts should not be mixed up. Normally, the area enclosed by the membership function and the abscissae, i.e. the axis referring to the elements of the basic set X, is not necessarily one. The interpretation differs as the value of a membership function $\mu_{\tilde{A}}(x)$ expresses the grade of membership of a certain element x to a fuzzy set $\tilde{A}$. In contrast, a density function specifies probabilities of subsets of a basic set Ω where $(\Omega, \mathcal{A}, P)$ is a probability space.

The difference between a crisp and a fuzzy set is also visualized in Figure 2.2. While it is easy to verify whether an element $x \in X$ belongs to the crisp set $C \subset X$ in Figure 2.2a an element x of a fuzzy set $\tilde{A}$ cannot be identified clearly in all cases in Figure 2.2b. The lighter the color in Figure 2.2b, the lower is the grade of membership. This comes along with Zadeh's interpretation who pointed out that fuzziness stands for diffuse, i.e. fuzzy, boundaries (cf. Kruse et al. 1995, p. vi).



(a) A crisp set $C \in \mathscr{P}(X)$. (b) A fuzzy set $\tilde{A} \in F\mathscr{P}(X)$.

Figure 2.2: Illustration of a crisp and a fuzzy subset of a basic set X.

In the following some characteristics of fuzzy sets are presented in order to describe them.

Definition 2.3 (Height of a fuzzy set)
Let $\tilde{A}$ be a fuzzy set. We call

$$\operatorname{hgt}(\tilde{A}) := \sup_{x \in X} \mu_{\tilde{A}}(x)$$

height of a fuzzy set $\tilde{A}$.

Definition 2.3 states that the height of a fuzzy set $\tilde{A}$ specifies the least upper bound of the membership function $\mu_{\tilde{A}}$. Consequently, it indicates the largest grade of membership of a fuzzy set. If the height is one, normal fuzzy sets are being considered which are defined in the following Definition 2.4.

Definition 2.4 (Normal fuzzy set and core)
Let $\tilde{A}$ be a fuzzy set. $\tilde{A}$ is called **normal** if $\operatorname{hgt}(\tilde{A}) = 1$. Moreover, the crisp set of elements having a degree of membership of one is refered to as the **core** of $\tilde{A}$, i.e. $\operatorname{core}(\tilde{A}) = \{x \in X \mid \mu_{\tilde{A}}(x) = 1\}$.

Remarks 2.5

a) Fuzzy sets $\tilde{A}$ of which the basic set is nonempty with a height strictly between zero and one, i.e. $0 < \operatorname{hgt}(\tilde{A}) < 1$, are called **unnormal**.

b) A nonempty fuzzy set $\tilde{A}$ can always be normalized by division of $\mu_{\tilde{A}}(x)$ by $\sup_{x \in X} \mu_{\tilde{A}}(x)$ for all $x \in X$.

c) Definition 2.4 states that a fuzzy set $\tilde{A}$ is normal if the core is nonempty.

For convenience sake we will always consider normal fuzzy sets if not explicitly stated otherwise. Referring to Definition 2.1, the concept of fuzzy sets and its associated membership function is a generalization of a set and its indicator function in the classical sense (cf. Zimmermann 2001, p. 14). While considering fuzzy sets it is possible that a set also comprises elements with grade of membership zero. Consequently, one is interested in those elements with a grade of membership unequal to zero.

Definition 2.6 (Support of a fuzzy set)
Let $\tilde{A}$ be a fuzzy set. The crisp set

$$\operatorname{supp}(\tilde{A}) := \{x \in X \mid \mu_{\tilde{A}}(x) > 0\}$$

is called **support** of $\tilde{A}$.

This issue can be generalized by the concept of an α-level set. So far only a representation with the help of a membership function has been considered. There exists an alternative representation via α-level sets, $\alpha \in \mathbb{R}$. We consider $X = \mathbb{R}$. Then, they denote unions of intervals $[a, b] \subseteq X$ in which the elements are assigned a grade of membership of at least α. This issue is sometimes referred to as *horizontal representation* of a fuzzy set (cf. Kruse et al. 1995, p. 16).[4] As we are mostly considering fuzzy sets on the basic set $X = \mathbb{R}$ we are restricting ourselves to that case if not explicitly stated otherwise.

Definition 2.7 ($\boldsymbol{\alpha}$-level set)
Let $\tilde{A}$ be a fuzzy set and $\alpha \in [0, 1]$. The set

$$\tilde{A}_\alpha := \{x \in X \mid \mu_{\tilde{A}}(x) \geq \alpha\}$$

is called **$\boldsymbol{\alpha}$-level set** or **$\boldsymbol{\alpha}$-level (cut)**. The set $\tilde{A}'_\alpha := \{x \in X \mid \mu_{\tilde{A}}(x) > \alpha\}$ is called **strong $\boldsymbol{\alpha}$-level set** or **strong $\boldsymbol{\alpha}$-level (cut)**.

The concept of an strong α-level set and support of a fuzzy set is visualized in Figure 2.3. All elements $x \in X$ having a membership grade exceeding the threshold α belong to the strong α-level set. A strong α-level set is a subset of the support where for a choice of $\alpha = 0$ the support is given, i.e. $\tilde{A}'_0 = \mathrm{supp}(\tilde{A})$. In fact, this concept allows us to form a family of crisp sets out of a fuzzy set if we build an α-level set for every $\alpha \in [0, 1]$. The notation α-level cut yields from the fact that the prescription for composing α-level sets divides up an originally fuzzy set into several crisp sets.

Moreover, every fuzzy set can be characterized by its α-cuts due to the following Proposition 2.8.

Proposition 2.8
Let $\tilde{A}$ be a fuzzy set over a basic set X and $\mu_{\tilde{A}}$ denotes the corresponding membership function. We then have

$$\mu_{\tilde{A}}(x) = \sup_{\alpha \in [0,1]} \left\{ \min\left(\alpha, \mathbf{1}_{\tilde{A}_\alpha}(x)\right)\right\}.$$

[4] Kruse et al. denote the representation via membership functions analogously as vertical representation.

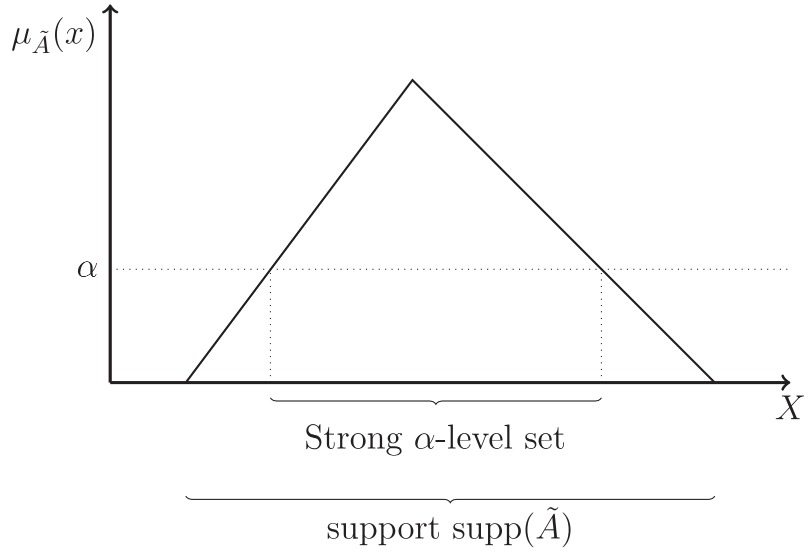


Figure 2.3: Strong α-level set $\tilde{A}'_\alpha$ and support $\text{supp}(\tilde{A})$ of a fuzzy set $\tilde{A}$.

Proof. Let $x \in X$ and $\alpha \in [0,1]$. It holds

$$\min\left(\alpha, \mathbf{1}_{\tilde{A}_\alpha}(x)\right) = \begin{cases} \alpha & \text{if } x \in \tilde{A}_\alpha \Leftrightarrow \mu_{\tilde{A}}(x) \geq \alpha \\ 0 & \text{if } x \notin \tilde{A}_\alpha \Leftrightarrow \mu_{\tilde{A}}(x) < \alpha \end{cases}.$$

We yield

$$\begin{aligned} \mu_{\tilde{A}}(x) &= \sup\{\alpha \mid \alpha \leq \mu_{\tilde{A}}(x)\} \\ &= \sup_{\alpha \in [0,1]} \left\{\min\left(\alpha, \mathbf{1}_{\tilde{A}_\alpha}(x)\right)\right\}. \end{aligned}$$

Another important concept in the theory of fuzzy sets is convexity. In crisp set theory convexity is defined with the help of the support whereas fuzzy set theory makes use of the membership function. This definition will be of use later on in order to define fuzzy numbers.

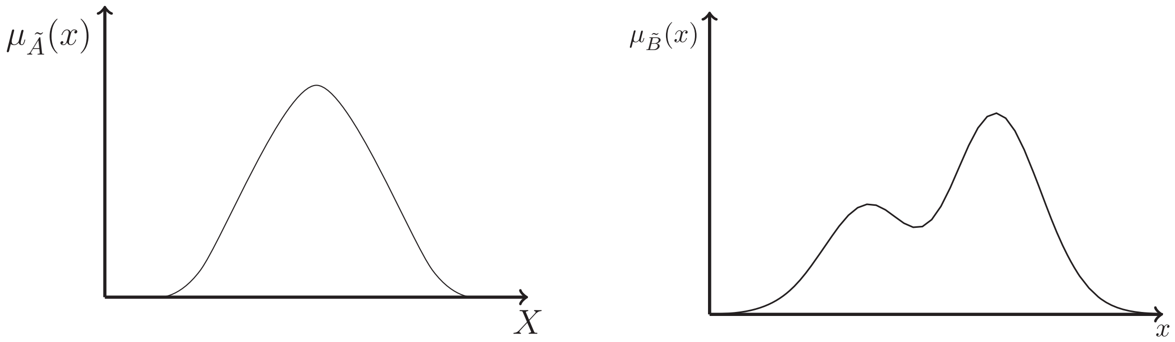
Definition 2.9 (Convex fuzzy set)
Let $\tilde{A}$ be a fuzzy set. A fuzzy set $\tilde{A}$ is called **convex** if

$$\mu_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geq \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)), \qquad x_1, x_2 \in X, \lambda \in [0,1].$$

Remark 2.10

We also speak of a convex fuzzy set if all α-level sets are convex (cf. Zimmermann 2001, p. 15). As α-level sets are crisp sets here the definition of convexity for crisp sets applies.

An illustration of a convex and a non-convex fuzzy set is given in Figure 2.4. In plain English a fuzzy set is convex if the graph of its membership function does not possess any "valleys".



(a) A convex fuzzy set $\tilde{A}$. (b) A non-convex fuzzy set $\tilde{B}$.

Figure 2.4: Convex and non-convex fuzzy sets

Like for crisp sets set operations can be also defined for fuzzy sets. Generally they are stated as defined in Definition 2.11 (cf. Zimmermann 2001, pp. 16ff.).

Definition 2.11

Let $\tilde{A}$ and $\tilde{B}$ be fuzzy sets and $\mu_{\tilde{A}}$ and $\mu_{\tilde{B}}$ denote their corresponding membership functions. We then define the following:

a) Two fuzzy sets $\tilde{A}$ and $\tilde{B}$ are said to be **equal** if

$$\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x) \quad \text{for all } x \in X.$$

b) The membership function $\mu_{\tilde{A}\cap\tilde{B}}$ of the **intersection** $\tilde{A} \cap \tilde{B}$ is defined for every $x \in X$ by

$$\mu_{\tilde{A}\cap\tilde{B}}(x) := \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}.$$

c) The membership function $\mu_{\tilde{A}\cup\tilde{B}}$ of the **union** $\tilde{A}\cup\tilde{B}$ is defined for every $x \in X$ by

$$\mu_{\tilde{A}\cup\tilde{B}}(x) := \max\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}.$$

d) The membership function $\mu_{\tilde{A}^C}$ of the complement $\tilde{A}^C$ of a normalized fuzzy set $\tilde{A}$ for every $x \in X$ is defined by

$$\mu_{\tilde{A}^C}(x) = 1 - \mu_{\tilde{A}}(x).$$

Remark 2.12
The definition of the membership function of the union and intersection is not intuitively at first hand. Bellman and Giertz came up with a justification for the definition with the help of the min- and max-operators (cf. Bellman and Giertz 1973).

2.1.3 The Extension Principle

The aim of this section is to present one of the most applied propositions in fuzzy set theory, the extension principle, which was firstly stated by Zadeh in a rather simple form (cf. Zadeh 1965, p. 346). The form which is used nowadays has been firstly shown in Zadeh (1975a,b,c).

The objective has been to transfer familiar operations like e.g. addition and multiplication to fuzzy sets. These will be used in Section 2.2. The definition will be stated in the sense of a generalization. Therefore, it is intended that the well-known operations are equivalent in the case of crisp sets. With the help of Definition 2.13 we can allow for fuzzy sets in a mapping instead of crisp elements.
We follow the presentations by Hanss (2005, pp. 44f.) and Zimmermann (2001, pp. 55f.).

Definition and Proposition 2.13 (Extension principle)
Let $X_1, \ldots, X_n, X$ be non-empty sets and

$$f:\ X_1 \times \ldots \times X_n \longrightarrow X$$

be a mapping. Let $\tilde{A}_1 \in F\mathscr{P}(X_1), \ldots \tilde{A}_n \in F\mathscr{P}(X_n)$ be fuzzy sets with corresponding membership functions

$$\mu_{\tilde{A}_1}:\ X_1 \longrightarrow [0,1], \ldots, \mu_{\tilde{A}_n}:\ X_n \longrightarrow [0,1].$$

We yield the membership function $\mu_{f(\tilde{A}_1,\ldots,\tilde{A}_n)}:\ X \longrightarrow [0,1]$ of a fuzzy set $f(\tilde{A}_1,\ldots,\tilde{A}_n) \in F\mathscr{P}(X)$ by

$$\mu_{f(\tilde{A}_1,\ldots,\tilde{A}_n)}(x) = \begin{cases} \sup\limits_{x=f(x_1,\ldots,x_n)} \min\left(\mu_{\tilde{A}_1}(x_1),\ldots,\mu_{\tilde{A}_n}(x_n)\right) & \text{if there exist } x_1,\ldots,x_n \\ & \text{with } x=f(x_1,\ldots,x_n) \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

and therefore a mapping

$$\tilde{f}:\ F\mathscr{P}(X_1)\times\ldots\times F\mathscr{P}(X_n) \longrightarrow F\mathscr{P}(X).$$

We say that $f:\ X_1\times\ldots\times X_n \longrightarrow X$ has been extended to $\tilde{f}:\ F\mathscr{P}(X_1)\times\ldots\times F\mathscr{P}(X_n) \longrightarrow F\mathscr{P}(X)$ by means of the **extension principle**.

Remarks 2.14

a) Every n-tuple $(x_1,\ldots,x_n) \in X_1\times\ldots\times X_n$ is mapped to an element $x \in X$ via the mapping f. First, the minimal grades of membership are determined component-wise. Then, the supremum of these grades of membership is taken as the grade of membership of the element x to the fuzzy set $f(\tilde{A}_1,\ldots,\tilde{A}_n)$.

b) If there exist no inverse images in (2.1) we use the common convention that $\sup\emptyset = 0$.

2.1.4 Relationship between Probability Theory and Fuzzy Set Theory

This section aims to highlight the differences between classical probability theory and fuzzy set theory. The prevailing discussion exists since the emergence of fuzzy set theory with the seminal paper of Lotfi A. Zadeh in 1965 (cf. Zadeh 1965). In the course of controversial discussions about the differences and similarities Kosko wonders: "Is uncertainty the same as randomness? If we are not sure about something, is it only up to chance?" (cf. Kosko 1990, p. 211). According to Forschner (1998) the conceptions differ widely: On the one hand probability theory is referred to as a special case of fuzzy set theory whereas on the other hand fuzziness is declared as "disguised" probability (cf. Forschner 1998, pp. 59f.). Nevertheless, Zadeh described fuzzy sets as "completely non statistical in nature" (cf. Zadeh 1965, p. 340) when introducing them. A small collection of criticism of fuzzy set theory is e.g. assembled in Dubois and Prade (2000, p. 3). Likewise, there are many publications of the advocators of fuzzy set theory whose publications provide critical comparisons of those two theories (cf. among others Zadeh 1995; Dubois and Prade 1994; Kosko 1990).[5]

There are differences considering the definitions as well as in the areas of application. The original aim of fuzzy set theory has been to model imprecision which can be hardly modeled with existing methods and vagueness especially in human language. Nowadays, a widely used method is e.g. "Computing with words" (CWW) (cf. Zadeh 1996).[6]

According to Zadeh (1995), fuzzy set theory does not only provide an opportunity to compute with words but also circumvents some drawbacks of probability theory which might not be useful in describing all kinds of imprecision and vagueness. Among others the following observations can be noted:[7]
Firstly, problems can occur when modeling a fuzzy event. In an insurance context there could be statements as e.g. "Tomorrow occurs a large claim." or "Many claims will be settled in the near future." which cannot be depicted properly with probability theory. Moreover, quantifiers as e.g. "many", "most", "several" or "few" cannot be modeled in an adequate way. A similar observation can be made for fuzzy occurrence probabilities. For instance, there exists no framework to map fuzzy event risks such as "likely", "unlikely" or "not very likely". Furthermore, there might be trouble if data is described with fuzzy expressions like e.g. "The portfolio consists of approximately 10,000 insured.". All in all Zadeh holds the view that probability theory cannot be properly applied in all areas in which dependencies between variables cannot be described precisely, probabilities are imprecise or human reasoning as well as perceptions and emotions of human beings are of great importance (cf. Zadeh 1995, p. 274). Zadeh concludes:

[5]In fact, Dubois and Prade (1994) and Dubois et al. (2000) also point out the links of those two theories especially regarding possibility distributions. The concept of possibility distributions will not be considered in the course of this work.
[6]A extensive background to fuzzy set theory and historical abstract can be found e.g. in Seising (2005).
[7]The not necessarily exhaustive list is based on Zadeh (1995, pp. 274f.).

> “The core of the position put forth [...] is that probability theory by itself is not sufficient and that it must be used in concert with fuzzy logic to enhance its effectiveness. In this perspective, probability theory and fuzzy logic are complementary rather than competitive.” (cf. Zadeh 1995, p. 271)

In the course of comparison of both methods two different approaches in probability theory need to be considered. On the one hand there is the axiomatic approach according to Kolmogorov, on the other hand there are interpretations regarding the description. The latter can be further subdivided into the frequentists’ view and the subjective view of probability theory.

The frequentists assume that an experiment is conducted arbitrarily many times. Then, the occurrence probability of an event is defined as the limit of its relative frequency. When contrasting fuzzy set theory and probability theory often only the frequentists’ view is taken into account. One much cited argument is that economic problems are usually irreproducible and, thus, probabilities – considered as relative frequencies – are not suitable (cf. Bosch 1993, p. 61).

In subjective probability theory a probability is a measure for the available information about an event (cf. Finetti 1981). As in Bayesian statistics subjective probabilities can be derived from expert opinions, market statistics, ... Hence, a clear separation as in the frequentist approach is not possible. However, Carlsson is of the opinion that “fuzziness is neither uncertainty in the sense of subjective probability – because it does not use its axioms” (cf. Carlsson 1984, p. 18).

Nowadays, probability theory is usually based on the axioms of Kolmogorov which were introduced in 1933 (cf. Kolmogorov 1933).

Definition 2.15 (Probability measure and probability space)
Let Ω be a set and $\mathcal{A}$ be a σ-algebra over Ω. Then a function $P : \mathcal{A} \to [0,1]$ is called probability measure P if the following axioms of Kolmogorov are fulfilled:

i) $P(A) \geq 0$ for all $A \in \mathcal{A}$

ii) $P(\Omega) = 1$

iii) Let $A_i \in \mathcal{A}$, $i \in \mathbb{N}$, be disjunct, i.e. $A_i \cap A_j = \emptyset$ for all $i \neq j$. Then

$$P\left(\cup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} P(A_i)$$

holds true (σ-additivity).

Then $(\Omega, \mathcal{A}, P)$ is called probability space.

When comparing probability measures and membership functions (cf. Definitions 2.1 and 2.15) some differences become apparent. While the domain of a membership function is the universe X, the domain of a probability measure is the σ-algebra $\mathcal{A}$. Furthermore, the values of a membership function do not need to sum up to one (see also Zimmermann 2001, p. 137). Hence, the axioms in fuzzy set theory are considered as not as strict as in probability theory. The same becomes apparent when comparing the characteristics of a probability density function and a normalized membership function. While the integral of the probability density function yields one neither the grades of membership do sum up to one nor the integral of the membership function yields one.

As presented in this section there are various differences between the underlying theory as well as the areas of applications. Thus, this work does not aim to present replacements of applications of probability theory and statistics in actuarial science and especially in claims reserving, but rather intends to show complementary suggestions to model fuzziness and imprecision in these fields. Zimmermann remarks that "Decision models might contain probabilistic as well as possibilistic components." (cf. Zimmermann 1983, p. 205) – a statement which cannot be limited to decision models. In fact, there are also attempts to bring together both concepts as e.g. fuzzy random variables or a cooperation under the guise of soft computing (cf. Zadeh 1995, p. 275).

Therefore we go in line with Dubois and Prade who state:

> "[...] we certainly do not believe that there exists a single theory that provides 'a complete and uniquely optimal means for solving problems and managing uncertainty'. Each concerned theory, whenever it is mathematically consistent (as is the case with probability theory, fuzzy set and possibility theory,[8] and several others), is only a convenient and appropriate means for solving a particular type of problem and grasping a particular facet of uncertainty." (cf. Dubois and Prade 1994, p. 20)

[8] Possibility theory is a subfield of fuzzy set theory.

2.2 Fuzzy Numbers

Fuzzy sets can be defined over various sets as there are no strict requirements. Of particular interest are fuzzy sets over the real numbers $\mathbb{R}$. There has been a demand for the investigation of fuzzy sets over $\mathbb{R}$ since in science or in practice there might be data which should be examined but is only approximately known or the data is sparse. The concept of a fuzzy number aims to make it possible to work with this kind of data by assigning an interval of numbers grades of membership. Therefore, every value is attached with a grade of plausibility.

Fuzzy numbers have been investigated since the mid 1970s. Overviews are e.g. given in Dubois and Prade (1993) and Hanss (2005). They are among others also considered in the textbooks Dubois and Prade (1980, pp. 26ff.), Zimmermann (2001, pp. 59ff.) and Bandemer and Gottwald (1993, pp. 29ff.).

2.2.1 Definitions and Types of Fuzzy Numbers

We follow the definition of a fuzzy number according to Hanss (cf. Hanss 2005, p. 45).

Definition 2.16 (Fuzzy number)
A fuzzy set $\tilde{a}$ over the real numbers $\mathbb{R}$ is called a **fuzzy number** (FN) if it satisfies the following conditions:

a) $\tilde{a}$ is convex.

b) There exists exactly one $x_0 \in \mathbb{R}$ for which $\mu_{\tilde{a}}(x_0) = 1$, i.e. the fuzzy set $\tilde{a}$ is normal and $\text{core}(\tilde{a}) = \{x_0\}$.

c) The membership function $\mu_{\tilde{a}}$ is piecewise continuous.

In the following we will call x_0 **modal value** or **mode** of a fuzzy number $\tilde{a}$. In the literature it is also referred to as *peak value*, *center value* or *mean value*. A FN can be interpreted as an imprecisely defined range of values. The support states an interval or a set of values which are plausible with the mode being the most plausible (cf. Dubois and Prade 1993, p. 118). It needs to be kept in mind that membership functions of FNs are often build solely by subjectivity as a source of information.

A general example of a fuzzy number can be seen in Figure 2.5 but it could be of a completely different shape as well.

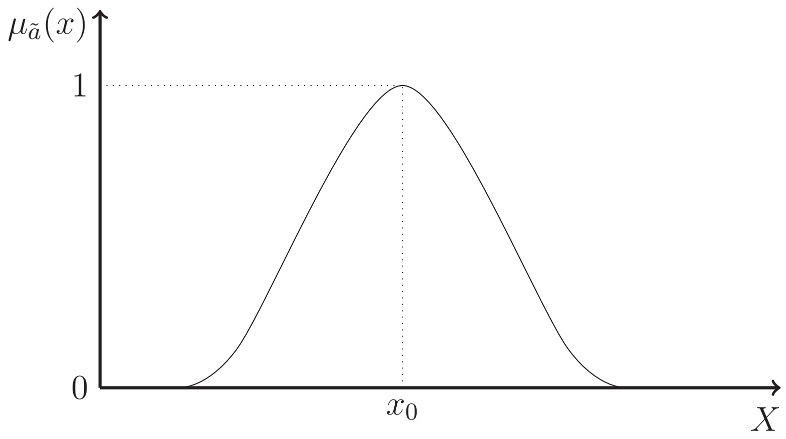


Figure 2.5: An example of a fuzzy number $\tilde{a}$.

A fuzzy number as defined in Definition 2.16 can possess the following characteristics:

Definition 2.17
Let $\tilde{a}$ be a fuzzy number with mode $x_0 \in \mathbb{R}$.

a) $\tilde{a}$ is called **symmetric** if

$$\mu_{\tilde{a}}(x_0 + x) = \mu_{\tilde{a}}(x_0 - x) \qquad \forall x \in \mathbb{R}.$$

b) $\tilde{a}$ is called **positive**, i.e. $\tilde{a} > 0$, if

$$\operatorname{supp}(\tilde{a}) \subseteq (0, \infty).$$

c) $\tilde{a}$ is called **negative**, i.e. $\tilde{a} < 0$, if

$$\operatorname{supp}(\tilde{a}) \subseteq (-\infty, 0).$$

d) $\tilde{a}$ is called a **(fuzzy) zero** if $0 \in \operatorname{supp}(\tilde{a})$.

According to Definition 2.16 there are no requirements towards the graph, resp. the shape, of a fuzzy number except for the convexity, unimodality and piecewise continuity.

In fact, there exist various types of fuzzy numbers which differ in shape but also in the way they can be explicitly calculated under arithmetical operations. Some are much easier to handle than others.
In the following we introduce several different types of fuzzy numbers and their membership functions. We follow the presentation of Hanss (cf. Hanss 2005, pp. 46ff.).

Some of the most used fuzzy numbers are triangular fuzzy numbers (TFNs). The name derives from the graph of the membership function. An example is visualized in Figure 2.6. One of their advantages is their simple representation as their membership functions are piecewise linear.

Definition 2.18 (Triangular fuzzy number (TFN))
Let $\tilde{a} \in F\mathscr{P}(\mathbb{R})$ and $\alpha_l, \alpha_r \geq 0$. A fuzzy number $\tilde{a} \in F\mathscr{P}(\mathbb{R})$ with mode $x_0 \in \mathbb{R}$ is called **triangular fuzzy number** if its membership function can be written in the form

$$\mu_{\tilde{a}}(x) = \begin{cases} 0 & \text{for } x \leq x_0 - \alpha_l \\ 1 + (x - x_0)/\alpha_l & \text{for } x_0 - \alpha_l < x < x_0 \\ 1 - (x - x_0)/\alpha_r & \text{for } x_0 \leq x < x_0 + \alpha_r \\ 0 & \text{for } x \geq x_0 + \alpha_r \end{cases}.$$

As before, x_0 is the mode of the fuzzy number and α_l and α_r are referred to as left and right spread, respectively. If the spreads are of equal length we will speak of a symmetric TFN (STFN) and we will denote a STFN by $\tilde{b} = (b, l_b)$.

Another type of fuzzy numbers is given by Gaussian fuzzy numbers. In contrast to the piecewise linear membership function leading to TFNs, the membership function in this case is characterized by a normalized Gaussian function. The functions describing the left and right spread can have the same “standard deviation” resulting in a symmetric fuzzy number but they do not need to be. In the latter case an asymmetric fuzzy number is on hand.

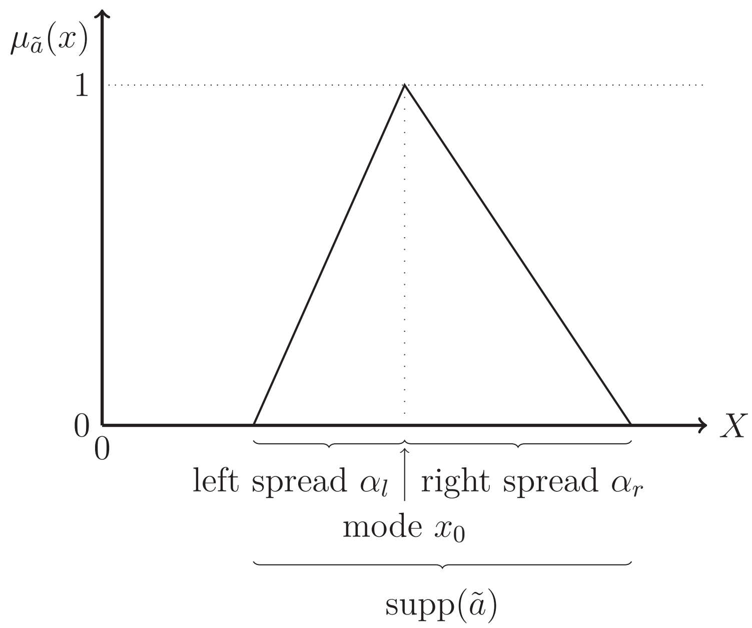


Figure 2.6: Example of a membership function of a triangular fuzzy number.

Definition 2.19 (Gaussian fuzzy number)
Let $\tilde{a} \in F\mathscr{P}(\mathbb{R})$ and $\sigma_l, \sigma_r \geq 0$. A fuzzy number $\tilde{a}$ with mode $x_0 \in \mathbb{R}$ is called **Gaussian fuzzy number** if its membership function can be written in the form

$$\mu_{\tilde{a}}(x) = \begin{cases} \exp\left(-\frac{(x-x_0)^2}{2\sigma_l^2}\right) & \text{for } x < x_0 \\ \exp\left(-\frac{(x-x_0)^2}{2\sigma_r^2}\right) & \text{for } x \geq x_0 \end{cases}.$$

As in the previous settings x_0 refers to the mode and σ_l and σ_r lead to the left and right spreads. Contrary to the situations of TFNs σ_l and σ_r do not denote the complete spread but correspond to the standard deviation of a Gaussian distribution. An example of a membership function of a Gaussian fuzzy number can be found in Figure 2.7. Along the lines of the definition of a fuzzy number with a piecewise linear membership function fuzzy numbers with a piecewise quadratic membership function are defined. Those are named quadratic fuzzy numbers.

Definition 2.20 (Quadratic fuzzy number)
Let $\tilde{a} \in F\mathscr{P}(\mathbb{R})$ and $\beta_l, \beta_r \in \mathbb{R}$. A fuzzy number $\tilde{a}$ with mode $x_0 \in \mathbb{R}$ is called **quadratic fuzzy number** if its membership function can be written in the form

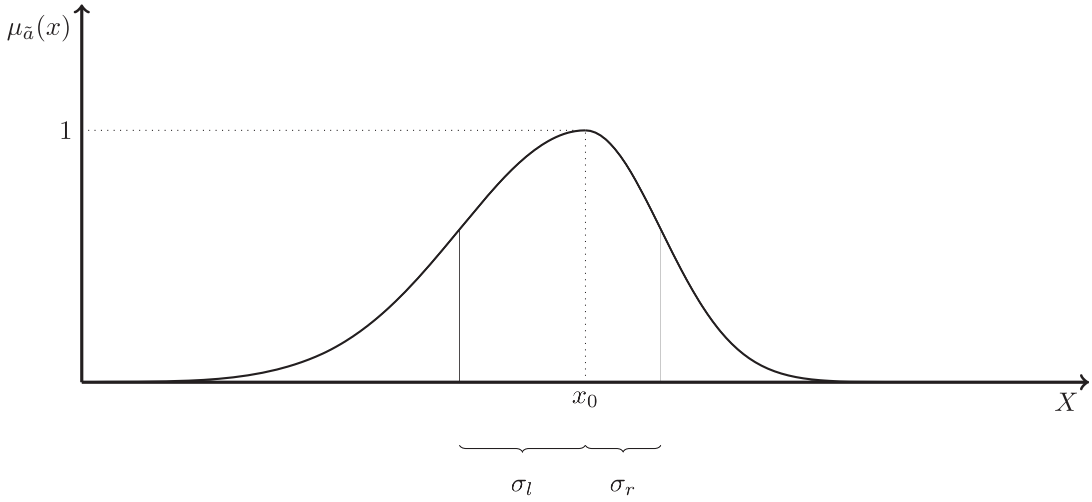


Figure 2.7: Example of a membership function of a Gaussian fuzzy number.

$$\mu_{\tilde{a}}(x) = \begin{cases} 0 & \text{for } x \leq x_0 - \beta_l \\ 1 - \frac{(x-x_0)^2}{\beta_l^2} & \text{for } x_0 - \beta_l < x < x_0 \\ 1 - \frac{(x-x_0)^2}{\beta_r^2} & \text{for } x_0 \leq x < x_0 + \beta_r \\ 0 & \text{for } x \geq x_0 + \beta_r \end{cases}.$$

An example of a fuzzy number with quadratic membership function is given in Figure 2.8.

As in the cases before, x_0 is the mode and β_l and β_r refer to the left and right spread. Membership functions of an exponential type lead to an exponential fuzzy number.

Definition 2.21 (Exponential fuzzy number)
Let $\tilde{a} \in F\mathscr{P}(\mathbb{R})$ and $\tau_l, \tau_r \in \mathbb{R}$. A fuzzy number $\tilde{a}$ with mode $x_0 \in \mathbb{R}$ is called **exponential fuzzy number** if its membership function can be written in the form

$$\mu_{\tilde{a}}(x) = \begin{cases} \exp(-(x-x_0)/\tau_l) & \text{for } x < x_0 \\ \exp(-(x-x_0)/\tau_r) & \text{for } x \geq x_0 \end{cases}.$$

The concept of an exponential fuzzy number is visualized in Figure 2.9.

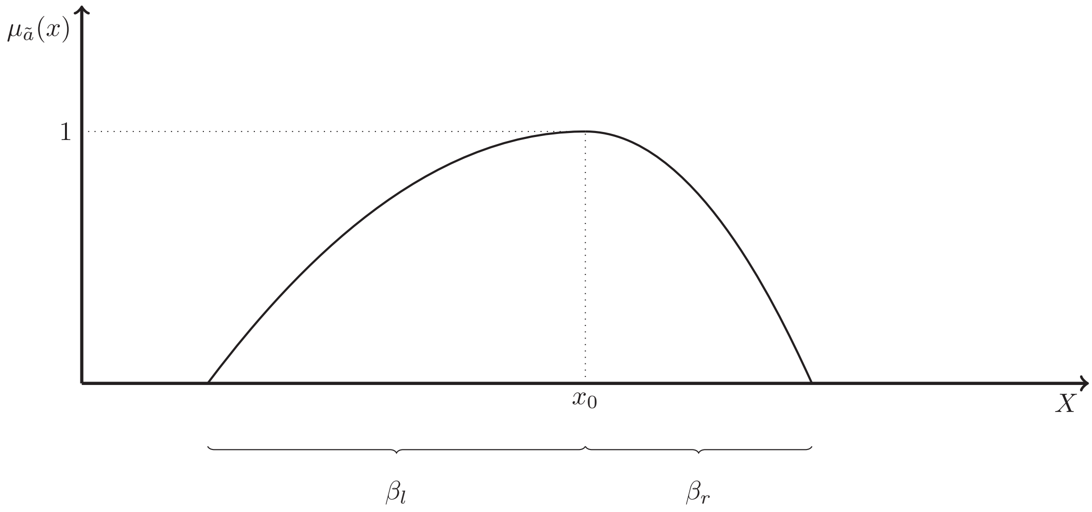


Figure 2.8: Example of a membership function of a quadratic fuzzy number.

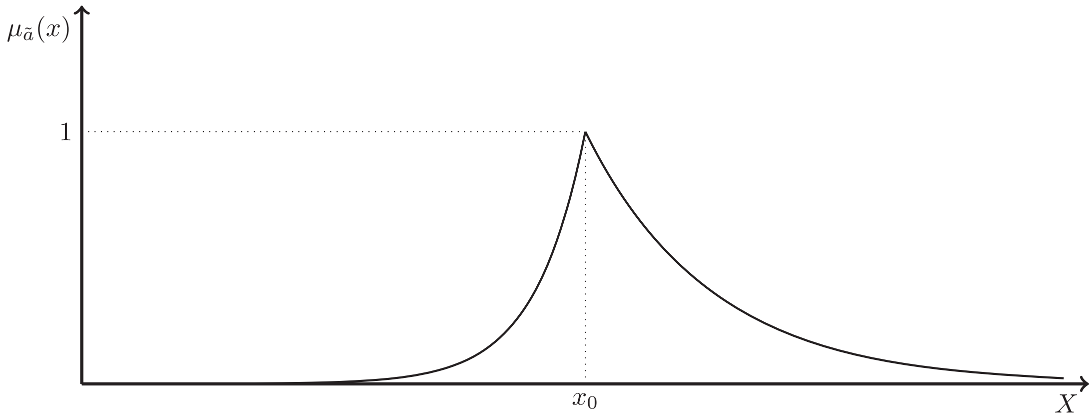


Figure 2.9: Example of a membership function of an exponential fuzzy number.

We introduced fuzzy sets as a generalization of the concept of crisp sets. In particular, the set of all crisp sets is a subset of the set of all fuzzy sets. Analogously, the set of crisp numbers in $\mathbb{R}$ can be seen as a subset of the set of all fuzzy numbers over $\mathbb{R}$. In fact, a fuzzy number $\tilde{a}$ over $\mathbb{R}$ given by the membership function

$$\mu_{\tilde{a}} = \begin{cases} 0 & \text{for } x < x_0 \\ 1 & \text{for } x = x_0 \\ 0 & \text{for } x > x_0 \end{cases} \tag{2.2}$$

refers to the crisp number x_0 (cf. Figure 2.10). When considering crisp numbers in the context of fuzzy numbers they are called **fuzzy singletons**.

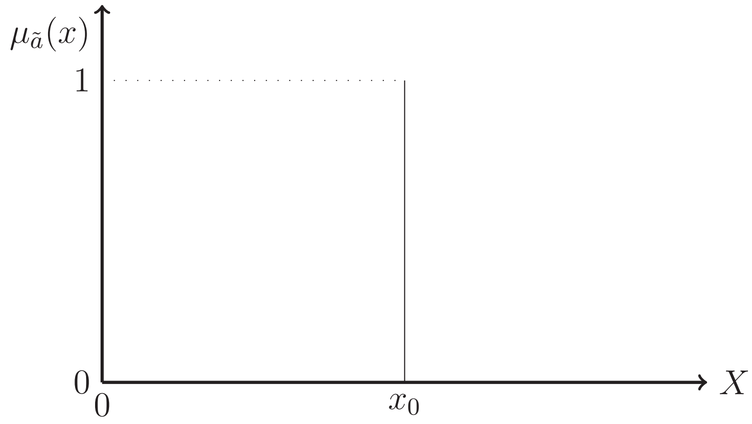


Figure 2.10: An example of a membership function of a fuzzy singleton.

A special type of fuzzy numbers has been presented by Dubois and Prade (cf. Dubois and Prade 1978). They introduced so-called *fuzzy numbers of LR-type* (or *LR-fuzzy numbers*) as a way to quickly compute operations on fuzzy numbers. They will be used in Section 2.2.2 and in particular in applications in Chapters 5 and 7. Their idea is to decompose a fuzzy number into two parts, i.e. one part to the left of the modal value and one to the right. For this definition functions describing the membership function on the left and right spread are needed.

Definition 2.22 (Reference function)
Let $L, R: \mathbb{R}_0^+ \longrightarrow [0,1]$ be mappings with the following properties:

a) $L(0) = R(0) = 1$.

b) L, R are strictly monotonically decreasing.

c) $$L(1) = 0 \text{ if } \min_{x\in[0,\infty)} L(x) = 0$$

$$\lim_{x\to\infty} L(x) = 0 \text{ if } L(x) > 0 \quad \text{for all } x \in \mathbb{R}_0^+$$

and

$$R(1) = 0 \text{ if } \min_{x\in[0,\infty)} R(x) = 0$$

$$\lim_{x\to\infty} R(x) = 0 \text{ if } R(x) > 0 \quad \text{for all } x \in \mathbb{R}_0^+.$$

Then L, R are called **reference functions**.

This leads to the following definition in which several types of fuzzy numbers can be comprised (cf. Hanss 2005, p. 54).

Definition 2.23 (Fuzzy number of LR-type)
Let L, R be reference functions. A fuzzy number $\tilde{a}$ is called **fuzzy number of LR-type** if the membership function can be represented as

$$\mu_{\tilde{a}}(x) = \begin{cases} L\left(\frac{a-x}{l_a}\right) & \text{for } x < a, l_a > 0 \\ R\left(\frac{x-a}{r_a}\right) & \text{for } x \geq a, r_a > 0 \end{cases}. \tag{2.3}$$

The mappings L and R are called **left** and **right reference** and we refer to $l_a > 0$ and $r_a > 0$ as the **left** and **right spread**. The value a is called **mode** or modal value.

Remarks 2.24

a) As an abbreviation of notation fuzzy numbers of LR-type $\tilde{a}$ will be denoted as

$$\tilde{a} := (a, l_a, r_a)_{L,R}\,.$$

b) If the reference functions L and R are identical, the fuzzy number of LR-type will be called semisymmetric. Moreover, if also the spreads are equal, i.e. $l_a = r_a$, it is designated as symmetric fuzzy number of LR-type.

c) In Definition 2.22 conditions a) and b) ensure that there exists exactly one modal value and that the resulting FN is convex. Since the functions L and R are mappings into the unit interval $[0, 1]$ the membership function can only take on values between zero and one. Condition c) requires that the membership values of the left and right spread approach zero (or at least approximately).

d) Many types of FNs can be conceived as FNs of LR-type, e.g. TFNs, Gaussian, quadratic and exponential FNs. Examples are given in Hanss (cf. 2005, p. 55).

Definition 2.25 (Opposite of a LR-fuzzy number)
Let $\tilde{a} = (a, l_a, r_a)_{L,R}$ be a fuzzy number of LR-type. The opposite $-\tilde{a}$ is defined as

$$-\tilde{a} := -(a, l_a, r_a)_{L,R} := (-a, r_a, l_a)_{R,L}\,.$$

Remark 2.26
In accordance with the definition of the negative of a crisp number we obtain the opposite of a fuzzy number of LR-type via mirroring along the grade of membership axis (y-axis).

Remark 2.27
When considering a function $\tilde{g}$ of two fuzzy numbers of LR-type the result $\tilde{g}\left(\tilde{a}, \tilde{b}\right)$ may not be obtained in all situations (cf. Dubois and Prade 1993, p. 131). Dubois and Prade (1993) propose to use Taylor's expansion in order to deduce an approximation which is given by a fuzzy number of LR-type (see also Dubois 1983, pp. 188f.).
Let $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a function and $\tilde{a} = (a, l_a, r_a)_{L,R}$ and $\tilde{b} = (b, l_b, r_b)_{L,R}$ two fuzzy numbers of LR-type. An approximation of $g(\tilde{a}, \tilde{b})$ is given by

$$g^* := \left(g(a,b), \frac{\partial g(a,b)}{\partial a} l_a + \frac{\partial g(a,b)}{\partial b} l_b, \frac{\partial g(a,b)}{\partial a} r_a + \frac{\partial g(a,b)}{\partial b} r_b\right).$$

2.2.2 Fuzzy Arithmetic

In the same way as we are interested in operations of crisp numbers there is also an interest in appropriate definitions for elementary operations such as addition, subtraction, multiplication (of two fuzzy numbers and with a scalar) and division for fuzzy numbers.[9] Therefore, the prior objective of fuzzy arithmetic is to specify the membership function of the resulting fuzzy number after an operation is applied. It turns out that not all operations are closed, i.e. that not all operations of FNs of the same type yield a FN of the identical type (e.g. operations of TFNs do not necessarily yield a TFN). Our objective is to determine the membership functions of the resulting fuzzy numbers for

[9] In the following we will stick to these notations for fuzzy operations and will use $+$, $-$, $\cdot$ and $\div$ in the crisp case.

$$\tilde{a}_1 \oplus \tilde{a}_2$$
$$\tilde{a}_1 \ominus \tilde{a}_2$$
$$\tilde{a}_1 \otimes \tilde{a}_2$$
$$\tilde{a}_1 \oslash \tilde{a}_2$$

where $\tilde{a}_1, \tilde{a}_2$ are fuzzy numbers and $\oplus$, $\ominus$, $\otimes$ and $\oslash$ stand for fuzzy addition, fuzzy subtraction, fuzzy multiplication and fuzzy division, respectively.

As a general approach we yield the membership function of the resulting fuzzy number with the help of the extension principle proposed by Zadeh (see Proposition 2.1). Fuzzy numbers are a special case of fuzzy sets and, thus, can be considered as such. Therefore, we use the corresponding membership functions as arguments in Equation (2.1), i.e. in the extension principle.
We denote by $O: \ F\mathscr{P}(\mathbb{R}) \times F\mathscr{P}(\mathbb{R}) \to F\mathscr{P}(\mathbb{R})$ a binary operation which can be one of the following: fuzzy addition, fuzzy subtraction, fuzzy multiplication or fuzzy division. The membership function of the resulting fuzzy number $\tilde{b} := O(\tilde{a}_1, \tilde{a}_2)$, $\tilde{a}_1, \tilde{a}_2 \in F\mathscr{P}(\mathbb{R})$, is achieved by the extension principle:

$$\mu_{\tilde{b}}(z) := \sup_{x_1, x_2 \in \mathbb{R}: O(x_1, x_2) = z} \min \{\mu_{\tilde{a}_1}(x_1), \mu_{\tilde{a}_2}(x_2)\} \tag{2.4}$$

In Definition 2.23 fuzzy numbers of LR-type have been introduced. They have been presented in Dubois and Prade (1978) in the late 1970's to develop easy-to-compute formulas for arithmetic operations.[10]

Proposition 2.28 (Addition of LR-fuzzy numbers)
Let $\tilde{a}_1 = (a_1, l_{a_1}, r_{a_1})_{L,R}$ and $\tilde{a}_2 = (a_2, l_{a_2}, r_{a_2})_{L,R}$ be two fuzzy numbers of the same LR-type. Let $a_1, a_2 \in \mathbb{R}, l_{a_1}, l_{a_2}, r_{a_1}, r_{a_2} > 0$. Their sum $\tilde{b}$ is given by

$$\tilde{b} = \tilde{a}_1 \oplus \tilde{a}_2 = (a_1 + a_2, l_{a_1} + l_{a_2}, r_{a_1} + r_{a_2})_{L,R}\,.$$

Proof. In case of the addition of two fuzzy numbers of the same LR-type the extension principle in Equation (2.4) gets:

[10] First results on operations of fuzzy numbers are shown in Dubois and Prade (1978). The work has been continued in Dubois and Prade (1979). A summary is provided by the authors in Dubois and Prade (1980) as well as by Hanss (2005).

$$\mu_{\tilde{a}_1 \oplus \tilde{a}_2}(z) = \sup_{x_1, x_2 \in \mathbb{R}: x_1 + x_2 = z} \min\{\mu_{\tilde{a}_1}(x_1), \mu_{\tilde{a}_2}(x_2)\}$$

We derive the membership function $\mu_{\tilde{a}_1 \oplus \tilde{a}_2}$ on the left hand side of the modal value first. For each $z \leq a_1 + a_2$ there exist $x_1^*, x_2^* \in \mathbb{R}$ with

$$x_1^* := a_1 - l_{a_1} \left(\frac{a_1 + a_2 - z}{l_{a_1} + l_{a_2}} \right) \leq a_1$$

and

$$x_2^* := a_2 - l_{a_2} \left(\frac{a_1 + a_2 - z}{l_{a_1} + l_{a_2}} \right) \leq a_2.$$

The corresponding values of their membership functions are equal (cf. Equation (2.3)):

$$\mu_{\tilde{a}_1}(x_1^*) = L \left(\frac{a_1 + a_2 - z}{l_{a_1} + l_{a_2}} \right) = \mu_{\tilde{a}_2}(x_2^*)$$

Hence, their minimum is given by

$$\min\{\mu_{\tilde{a}_1}(x_1^*), \mu_{\tilde{a}_2}(x_2^*)\} = L \left(\frac{a_1 + a_2 - z}{l_{a_1} + l_{a_2}} \right) =: \mu^*(z).$$

For all $\varepsilon \geq 0$ the inequation

$$\mu_{\tilde{a}_2}(x_2^* - \varepsilon) \leq \mu_{\tilde{a}_2}(x_2^*) = \mu^*(z)$$

and for all $\varepsilon \leq 0$ the inequation

$$\mu_{\tilde{a}_1}(x_1^* + \varepsilon) \leq \mu_{\tilde{a}_1}(x_1^*) = \mu^*(z)$$

hold true since the membership functions are increasing for all $x_1 \leq a_1$ and $x_2 \leq a_2$. Therefore, the minimum is bounded by $\mu^*(z)$ for all $\varepsilon \in \mathbb{R}$:

$$\min\{\mu_{\tilde{a}_1}(x_1^* + \varepsilon), \mu_{\tilde{a}_2}(x_2^* - \varepsilon)\} \leq \mu^*(z)$$

Hence, the extension principle yields for the left hand side of the membership function

$$\mu_{\tilde{a}_1 \oplus \tilde{a}_2}(z) = \sup_{x_1, x_2 \in \mathbb{R}: x_1 + x_2 = z} \min\{\mu_{\tilde{a}_1}(x_1), \mu_{\tilde{a}_2}(x_2)\} = \mu^*(z) = L \left(\frac{a_1 + a_2 - z}{l_{a_1} + l_{a_2}} \right)$$

for all $z \leq a_1 + a_2$. The proof for the right hand side of the membership function $\mu_{\tilde{a}_1 \oplus \tilde{a}_2}$ is done analogously and we yield

$$\mu_{\tilde{a}_1 \oplus \tilde{a}_2} = \mu^*(z) = R\left(\frac{z-(a_1+a_2)}{r_{a_1}+r_{a_2}}\right)$$

for all $z \geq a_1 + a_2$. □

Remarks 2.29

a) The sum of two fuzzy numbers of LR-type is again a fuzzy number of LR-type with the same reference functions L and R. The modal value is given by $a_1 + a_2$ and the left and right spreads by $l_{a_1} + l_{a_2}$ and $r_{a_1} + r_{a_2}$, respectively.

b) The addition of fuzzy numbers of LR-type is commutative and associative since for arbitrary $\tilde{a}_1 := (a_1, l_{a_1}, r_{a_1})_{L,R}$, $\tilde{a}_2 := (a_2, l_{a_2}, r_{a_2})_{L,R}$, $\tilde{a}_3 := (a_3, l_{a_3}, r_{a_3})_{L,R}$ of LR-type with $a_1, a_2, a_3 \in \mathbb{R}$ and $l_{a_1}, l_{a_2}, l_{a_3}, r_{a_1}, r_{a_2}, r_{a_3} > 0$ holds:

$$\begin{aligned}
\tilde{a}_1 \oplus \tilde{a}_2 &= (a_1 + a_2, l_{a_1} + l_{a_2}, r_{a_1} + r_{a_2})_{L,R} \\
&= (a_2 + a_1, l_{a_2} + l_{a_1}, r_{a_2} + r_{a_1})_{L,R} = \tilde{a}_2 \oplus \tilde{a}_1 \\
(\tilde{a}_1 \oplus \tilde{a}_2) \oplus \tilde{a}_3 &= (a_1 + a_2, l_{a_1} + l_{a_2}, r_{a_1} + r_{a_2})_{L,R} \oplus (a_3, l_{a_3}, r_{a_3})_{L,R} \\
&= ((a_1 + a_2) + a_3, (l_{a_1} + l_{a_2}) + l_{a_3}, (r_{a_1} + r_{a_2}) + r_{a_3})_{L,R} \\
&= (a_1 + (a_2 + a_3), l_{a_1} + (l_{a_2} + l_{a_3}), r_{a_1} + (r_{a_2} + r_{a_3}))_{L,R} \\
&= \tilde{a}_1 \oplus (\tilde{a}_2 \oplus \tilde{a}_3)
\end{aligned}$$

The properties of the fuzzy addition are inherited by the properties of the addition of real numbers $\mathbb{R}$.

c) The set of all fuzzy numbers of LR-type in conjunction with the fuzzy addition $(F\mathscr{P}(\mathbb{R}))_{L,R}, \oplus)$ is not an algebraic group. The neutral element is given by the fuzzy singleton $\tilde{0} := (0,0,0)_{L,R}$ so that for every $\tilde{a} \in F\mathscr{P}(\mathbb{R}))_{L,R}$ we have $\tilde{a} \oplus \tilde{0} = \tilde{a}$. However, there exists no inverse element for a given element $\tilde{a} \in F\mathscr{P}(\mathbb{R})_{L,R}$. If looking for an element $\tilde{a}' \in F\mathscr{P}(\mathbb{R})_{L,R}$ which satisfies $\tilde{a} \oplus \tilde{a}' = \tilde{0}$ the resulting system of equations would lead to negative spreads.

Proposition 2.30 (Subtraction of LR-fuzzy numbers)

Let $\tilde{a}_1 = (a_1, l_{a_1}, r_{a_1})_{L,R}$, $\tilde{a}_2 = (a_2, l_{a_2}, r_{a_2})_{R,L}$ be two fuzzy numbers where the functions l and R are identical. Let $a_1, a_2 \in \mathbb{R}$ and $l_{a_1}, l_{a_2}, r_{a_1}, r_{a_2} > 0$. Their difference is given by

$$\tilde{b} = \tilde{a}_1 \ominus \tilde{a}_2 = (a_1, l_{a_1}, r_{a_1})_{L,R} \ominus (a_2, l_{a_2}, r_{a_2})_{R,L} = (a_1 - a_2, l_{a_1} + r_{a_2}, r_{a_1} + l_{a_2})_{L,R} \,.$$

Proof. The proof uses the same technique as in Proposition 2.28. We yield the membership function of the subtraction of two fuzzy numbers with the help of the extension principle as in Equation (2.4):

$$\mu_{\tilde{a}\ominus\tilde{a}_2}(z) = \sup_{x_1,x_2\in\mathbb{R}:x_1-x_2=z} \min\{\mu_{\tilde{a}_1}(x_1), \mu_{\tilde{a}_2}(x_2)\}$$

We derive the left hand side of the membership function first. Let $z \in \mathbb{R}$. For $z \leq a_1 - a_2$ there exist x_1^* and x_2^* with

$$x_1^* = a_1 - l_{a_1}\left(\frac{a_1 - a_2 - z}{l_{a_1} + r_{a_2}}\right) \leq a_1$$

and

$$x_2^* = a_2 + r_{a_2}\left(\frac{a_1 - a_2 - z}{l_{a_1} + r_{a_2}}\right) \geq a_2.$$

The corresponding values of their membership functions are identical since

$$\mu_{\tilde{a}_1}(x_1^*) = L\left(\frac{a_1 - a_2 - z}{l_{a_1} + r_{a_2}}\right) = \mu_{\tilde{a}_2}(x_2^*).$$

Consequently the minimum is given by

$$\min\{\mu_{\tilde{a}_1}(x_1), \mu_{\tilde{a}_2}(x_2)\} = L\left(\frac{a_1 - a_2 - z}{l_{a_1} + r_{a_2}}\right) =: \mu_*(z).$$

For all $\varepsilon \geq 0$ the inequality

$$\mu_{\tilde{a}_2}(x_2^* + \varepsilon) \leq \mu_{\tilde{a}_2}(x_2^*) = \mu_*(z)$$

holds true since the membership function $\mu_{\tilde{a}_2}$ is decreasing for all $x_2 > a_2$. For all $\varepsilon \leq 0$ the inequality

$$\mu_{\tilde{a}_1}(x_1^* + \varepsilon) \leq \mu_{\tilde{a}_1}(x_1^*) = \mu_*(z)$$

holds true since the membership function $\mu_{\tilde{a}_1}$ is increasing for all $x_1 \leq a_1$. Hence, the minimum of the considered values of the membership functions is bounded by $\mu_*(z)$ in the following way:

$$\min\{\mu_{\tilde{a}_1}(x_1^* + \varepsilon), \mu_{\tilde{a}_2}(x_2^* + \varepsilon)\} \leq \mu_*(z)$$

Therefore, we yield for the membership function with the extension principle

$$\mu_{\tilde{a}_1 \oplus \tilde{a}_2}(z) = \sup_{x_1, x_2 \in \mathbb{R}: x_1 + x_2 = z} \min\{\mu_{\tilde{a}_1}(x_1), \mu_{\tilde{a}_2}(x_2)\} = L\left(\frac{a_1 - a_2 - z}{l_{a_1} + r_{a_2}}\right)$$

for all $z \leq a_1 - a_2$. The proof for the right hand side of the membership function of the subtraction of two fuzzy numbers is done analogously. $\square$

A formula for the multiplication of two fuzzy numbers of LR-type cannot be derived in the same way as for the addition. By doing so it emerges that the multiplication is not closed, i.e. the product of two fuzzy numbers of LR-type does not need to be of LR-type as well.
Let $\tilde{a}_1 = (a_1, l_{a_1}, r_{a_1})_{L,R}$ and $\tilde{a}_2 = (a_2, l_{a_2}, r_{a_2})_{L,R}$ be two fuzzy numbers of LR-type. Note that again we consider two fuzzy numbers with the same reference functions. We take on the assumption that both $\tilde{a}_1$ and $\tilde{a}_2$ are positive. The membership functions of the left references are given by

$$\mu_{L,\tilde{a}_1}(x_1) = L\left(\frac{a_1 - x_1}{l_{a_1}}\right) \qquad \text{and} \qquad \mu_{L,\tilde{a}_2}(x_2) = L\left(\frac{a_2 - x_2}{l_{a_2}}\right),$$
$$x_1, x_2 \in \mathbb{R}, x_1 \leq a_1, x_2 \leq a_2.$$

Let again $\mu^* \in [0,1]$ be arbitrary but fixed. The grade of membership μ^* is attained by the elements x_1^* and x_2^* with

$$x_1^* = a_1 - l_{a_1} L^{-1}(\mu^*) \qquad \text{and} \qquad x_2^* = a_2 - l_{a_2} L^{-1}(\mu^*).$$

As a consequence the product of these two values

$$z^* := x_1^* x_2^* = a_1 a_2 - (a_1 l_{a_2} + a_2 l_{a_1}) L^{-1}(\mu^*) + l_{a_1} l_{a_2} (L^{-1}(\mu^*))^2 \tag{2.5}$$

takes on the same grade of membership. Therefore, it results in a quadratic equation which possible solutions are given by

$$\left(L^{-1}(\mu^*)\right)_{1,2} = \frac{a_1 l_{a_2} + a_2 l_{a_1}}{2 l_{a_1} l_{a_2}} \pm \sqrt{\frac{(a_1 l_{a_2} + a_2 l_{a_1})^2}{4 l_{a_1}^2 l_{a_2}^2} - \frac{a_1 a_2 - z^*}{l_{a_1} l_{a_2}}}.$$

Due to the quadratic term the operation is not closed and normally no LR-fuzzy number results. That is the reason why Dubois and Prade proposed to omit the quadratic term (cf. Dubois and Prade 1980, p. 55). As the spreads are often small compared to the modal values this is a comprehensible approach. By neglecting the quadratic term in (2.5) we yield:

$$\begin{aligned} z^* &\approx a_1 a_2 - (a_1 l_{a_2} + a_2 l_{a_1}) L^{-1}(\mu^*) \\ \Leftrightarrow \mu^* &= L\left(\frac{a_1 a_2 - z^*}{a_1 l_{a_2} + a_2 l_{a_1}}\right) \end{aligned}$$

With the same arguments the right reference can be derived. Therefore, we get an approximation for the multiplication of two fuzzy numbers of LR-type by

$$(a_1, l_{a_1}, l_{a_2})_{L,R} \otimes (a_2, l_{a_2}, r_{a_2})_{L,R} \approx (a_1 a_2, a_1 l_{a_2} + a_2 l_{a_1}, a_1 r_{a_2} + a_2 r_{a_1})_{L,R}\,. \tag{2.6}$$

This formula is referred to as **tangent approximation** (Hanss 2005, p. 58).

Another approximation is given by the so-called *secant approximation*. If the spreads are not comparatively small to the modal values omitting the quadratic term is not reasonable. In this case a solution might be given by approximating the quadratic term by a linear one, i.e. replacing $(L^{-1}(\mu^*))^2$ by $L^{-1}(\mu^*)$. Then, the reference function for the left side can be derived by:

$$\begin{aligned} z^* &= a_1 a_2 - (a_1 l_{a_2} + a_2 l_{a_1}) L^{-1}(\mu^*) + l_{a_1} l_{a_2} L^{-1}(\mu^*) \\ \Leftrightarrow \mu^* &= L\left(\frac{z^* - a_1 a_2}{-a_1 l_{a_2} - a_2 l_{a_1} + l_{a_1} l_{a_2}}\right) = L\left(\frac{a_1 a_2 - z^*}{a_1 l_{a_2} + a_2 l_{a_1} - l_{a_1} l_{a_2}}\right) \end{aligned}$$

With the same arguments for the right reference another approximation is deduced such that we yield the following approximation which is referred to as **secant approximation**:

$$\begin{aligned} &(a_1, l_{a_1}, r_{a_1})_{L,R} \otimes (a_2, l_{a_2}, r_{a_2})_{L,R} \\ &\approx (a_1 a_2, a_1 l_{a_2} + a_2 l_{a_1} - l_{a_1} l_{a_2}, a_1 r_{a_2} + a_2 r_{a_1} + r_{a_1} r_{a_2})_{L,R} \end{aligned} \tag{2.7}$$

Tangent approximation	
Factors	Approximated product
$(a_1, l_{a_1}, r_{a_1})_{L,R} > 0$, $(a_2, l_{a_2}, r_{a_2})_{L,R} > 0$	$(a_1a_2, a_1l_{a_2} + a_2l_{a_1}, a_1r_{a_2} + a_2r_{a_1})_{L,R}$
$(a_1, l_{a_1}, r_{a_1})_{L,R} < 0$, $(a_2, l_{a_2}, r_{a_2})_{L,R} < 0$	$(a_1a_2, -a_1r_{a_2} - a_2r_{a_1}, -a_1l_{a_2} - a_2l_{a_1})_{R,L}$
$(a_1, l_{a_1}, r_{a_1})_{R,L} < 0$, $(a_2, l_{a_2}, r_{a_2})_{L,R} > 0$	$(a_1a_2, a_2l_{a_1} - a_1r_{a_2}, a_2r_{a_1} - a_1l_{a_2})_{R,L}$
Secant approximation	
Factors	Approximated product
$(a_1, l_{a_1}, r_{a_1})_{L,R} > 0$, $(a_2, l_{a_2}, r_{a_2})_{L,R} > 0$	$(a_1a_2, a_1l_{a_2} + a_2l_{a_1} - l_{a_1}l_{a_2}, a_1r_{a_2} + a_2r_{a_1} + r_{a_1}r_{a_2})_{L,R}$
$(a_1, l_{a_1}, r_{a_1})_{L,R} < 0$, $(a_2, l_{a_2}, r_{a_2})_{L,R} < 0$	$(a_1a_2, -a_1r_{a_2} - a_2r_{a_1} - r_{a_1}r_{a_2}, -a_1l_{a_2} - a_2l_{a_1} + l_{a_1}l_{a_2})_{R,L}$
$(a_1, l_{a_1}, r_{a_1})_{R,L} < 0$, $(a_2, l_{a_2}, r_{a_2})_{L,R} > 0$	$(a_1a_2, a_2l_{a_1} - a_1r_{a_2} + l_{a_1}r_{a_2}, a_2r_{a_1} - a_1l_{a_2} - r_{a_1}l_{a_2})_{R,L}$

Table 2.1: Overview of formulas for tangent and secant approximation (adapted from Hanss 2005, p. 60).

So far only approximation formulas for the case of two positive fuzzy numbers have been deduced. Table 2.1 gives an overview of the formulas for other cases.

The following observations can be remarked.

Remarks 2.31

a) With the multiplication of two fuzzy numbers of LR-type the fuzziness increases, i.e. the spreads of the resulting product broaden.

b) As for the addition the set of all fuzzy numbers of LR-type equipped with the fuzzy multiplication $(F\mathscr{P}(\mathbb{R})_{L,R}, \otimes)$ is not a group.

c) Wagenknecht et al. (2001) investigated upper and lower bounds for the exact multiplication of two non-negative TFNs $\tilde{a} := (a, l_a, r_a)_{L,R}$ and $\tilde{b} := (b, l_b, r_b)_{L,R}$. The derived lower bound is given by $(\tilde{a} \otimes \tilde{b})^{\text{lower}} := (ab, al_b + bl_a - l_al_b, ar_b + br_a)_{L,R}$ and the upper bound can be stated as $(\tilde{a} \otimes \tilde{b})^{\text{upper}} := (ab, al_b + bl_a, ar_b + br_a + r_ar_b)_{L,R}$. The left spread of the lower bound and the right spread of the upper bound represent the left and right spread of the secant approximation, respectively. In the following Chapters 5, 6 and 7 fuzzy multiplication (with secant approximation) will be used to derive predictions of reserves. Then, the chosen fuzzy multiplication is rather conservative. In this context conservative means that the resulting reserves are (slightly) overestimated.

Example 2.32

Let $\tilde{a} = (3,1,2)_{L,R}$ and $\tilde{b} = (2,2,1)_{L,R}$ be two TFNs. The aim of this example is to derive the product of these TFNs $\tilde{c} = \tilde{a} \otimes \tilde{b}$ with the help of tangent approximation as defined in Equation (2.6) and secant approximation as defined in Equation (2.7). Moreover, the membership function for the exact multiplication is deduced. All three cases are visualized in Figure 2.11.

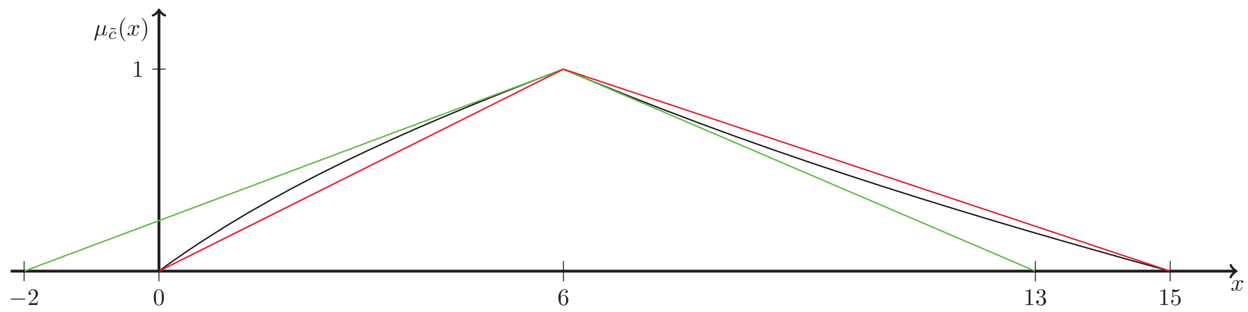


Figure 2.11: Comparison of tangent approximation (green line), secant approximation (red line) as well as the exact product (black line) of TFNs $\tilde{a} = (3,1,2)_{L,R}$ and $\tilde{b} = (2,2,1)_{L,R}$.

The black line is showing the resulting membership function for the exact multiplication. The red line refers to the secant approximation and the green line stands for tangent approximation. One observes that the support for the product resulting from exact multiplication and the secant approximation is identical. Moreover, the red line for the right spread is slightly above the black line which visualizes the fact mentioned in Remark 2.31c). The support of the product deduced by tangent approximation and the exact multiplication differ. Moreover, there are greater deviations in the "outer" spreads in this example.

The remaining part of the section is devoted to the definition of the quotient of two FNs of LR-type. In order to define the quotient of two fuzzy numbers of LR-type $\tilde{a}$ and $\tilde{b}$ one makes use of the following relation using the above defined multiplication

$$\tilde{a} \oslash \tilde{b} = \tilde{a} \otimes \tilde{b}^{-1}. \tag{2.8}$$

Hence, only the definition of the inverse is needed. According to Hanss (2005, pp. 60f.) it can be stated in the following way.

Proposition 2.33 (Inverse of a LR-Fuzzy Number)
*Let $\tilde{a} = (a, l_a, r_a)_{L,R}$ be a fuzzy number of LR-type which is either positive or negative. Then, the **fuzzy inverse** can be approximated by either*

$$\tilde{a}_{tan}^{-1} \approx \left(\frac{1}{a}, \frac{r_a}{a^2}, \frac{l_a}{a^2}\right)_{R,L} \qquad \textit{tangent approximation}$$

or

$$\tilde{a}_{sec}^{-1} \approx \left(\frac{1}{a}, \frac{r_a}{a(a+r_a)}, \frac{l_a}{a(a-l_a)}\right)_{R,L} \qquad \textit{secant approximation.} \tag{2.9}$$

Remark 2.34

a) The approximations given in Proposition 2.33 are more precise, the closer the considered value is to the mode of the FN (cf. Dubois and Prade 1979, p. 341).

b) As for the exact multiplication, Wagenknecht et al. (2001) also investigated upper and lower bounds for the exact inverse of a TFN $\tilde{a} = (a, l_a, r_a)_{L,R}$ (see Remark 2.31c)). These are given by $(\tilde{a}^{-1})^{\text{lower}} = \left(\frac{1}{a}, \frac{r_a}{a(a+r_a)}, \frac{l_a}{a^2}\right)_{R,L}$ and $(\tilde{a}^{-1})^{\text{upper}} = \left(\frac{1}{a}, \frac{r_a}{a^2}, \frac{l_a}{a(a-l_a)}\right)_{R,L}$. Comparing to Proposition 2.33 the left spread of the lower bound and the right spread of the upper bound refer to the left and right spread of the fuzzy inverse with secant approximation. With the same arguments as in Remark 2.31c) the approximation of the inverse is rather conservative.

Using the relation in Equation (2.8) the quotient of two fuzzy numbers of LR-type can be deduced with the help of Proposition 2.33. An overview of formulas for the fuzzy quotient is e.g. given in Hanss (2005, p. 61).

In the following Chapters 5, 6 and 7 we are dealing with TFNs which are by definition of LR-type. As the reference functions are all of the same type in that case, i.e. linear functions, we will drop the subscript "LR" and "RL", respectively, of the FNs in the following and write $\tilde{a} = (a, l_a, r_a)$ for a TFN $\tilde{a}$. Moreover, for the sake of readability we will omit the approximation symbol "≈" in the fuzzy multiplication and for the fuzzy inverse and write "=" instead.

2.2.3 Defuzzification Methods

The use of fuzzy sets and methods for modeling and describing issues yields fuzzy quantities. In fact, many procedures need a single crisp outcome and cannot handle e.g. a linguistic value or vague magnitude. For instance, in control theory even though the basis for decision is fuzzy a controller requires a crisp input. As a further example non-life actuaries have to set up a crisp reserve when using fuzzy methods as the balance sheet asks for crisp values. Therefore, there is a need for defuzzification methods which reduce a fuzzy set to a crisp value.

Definition 2.35 (Defuzzification operator)
Let X be a basic set. The mapping

$$D: \quad F\mathscr{P}(X) \to X, \quad A \mapsto D(A)$$

is called **defuzzification operator**.

Since we are mainly dealing with fuzzy numbers which are fuzzy sets over the real numbers $\mathbb{R}$ in the present work we are restricting the underlying set in Definition 2.35 to the case of $X = \mathbb{R}$.

In literature various defuzzification procedures are presented. A helpful overview is given in Leekwijk and Kerre (1999), Runkler and Glesner (1993) and Zimmermann (2001, pp. 232 – 239). According to Leekwijk and Kerre (1999) the techniques can be generally classified into three main groups: maxima, distribution and area methods.[11] The maxima procedures take one of the elements with the highest grade of membership as they follow the principle that a high grade of membership indicates that an element is suitable. Distribution methods consider the membership function as a density function as in probability theory and then calculate the expected value. Some of them also take

[11] Leekwijk and Kerre (1999) do not only classify many of the used defuzzification methods but also present a catalog of desirable properties the procedures aim to possess. Subsequently, they examine which characteristics the stated defuzzification techniques have. A similar approach is done in Runkler and Glesner (1993) whereas they postulate a little different properties.

into account different weighting functions. Whereas the area methods as the third class consider the area under the membership function to define the defuzzified value. Due to the variety of defuzzification methods we only state those used in the present work and alternatively refer to the literature.

One procedure using a weighted average is given by the center of gravity.

Definition 2.36 (Center of Gravity)
Let $\tilde{X} \in F\mathscr{P}(\mathbb{R})$ and $X \subseteq \mathbb{R}$. The **center of gravity** is given by

$$\mathrm{COG}(\tilde{X}) = \frac{\int\limits_{\mathrm{supp}(\tilde{X})} x\mu_{\tilde{X}}(x)\,dx}{\int\limits_{\mathrm{supp}(\tilde{X})} \mu_{\tilde{X}}(x)\,dx}$$

or

$$\mathrm{COG}(\tilde{X}) = \frac{\sum_{i=1}^{n} x_i\mu_{\tilde{X}}(x_i)}{\sum_{i=1}^{n} \mu_{\tilde{X}}(x_i)} \text{ in the finite case, i.e. if } X = \{x_1, \dots, x_n\}.$$

Referring to a statistical point of view the center of gravity complies with the expected value. From a geometrical view point it calculates the center of mass (cf. Zimmermann 2001, p. 234).

Remarks 2.37

a) For a triangular fuzzy number $\tilde{a} = (a, l_a, r_a)$ with a membership function as given in Definition 2.18 $\mathrm{COG}(\tilde{a})$ can be calculated as

$$\mathrm{COG}(\tilde{a}) = a + \frac{-\frac{1}{6}l_a^2 + \frac{1}{6}r_a^2}{\frac{1}{2}l_a + \frac{1}{2}r_a}.$$

Thus, for symmetric triangular fuzzy numbers the COG is the mode a.

b) The center of gravity does not necessarily yield a result which is an element of the support of the fuzzy set. However, when considering fuzzy numbers which are normal and convex fuzzy sets we can omit this problem: In that case, the support and all α-cuts are compact intervals.

c) The center of gravity has some desirable properties in addition to b) (cf. Leekwijk and Kerre 1999, pp. 173, 175–176; Zimmermann 2001, pp. 237f.):

  - For a fuzzy singleton it yields the only element with positive grade of membership.

– When the elements are translated in x-direction by $b \in \mathbb{R}$, i.e. $\tilde{b} = \tilde{a} \oplus (b, 0, 0)$, the relative position of the defuzzified value should stay the same. The corresponding membership function $\mu_{\tilde{b}} : \mathbb{R} \longrightarrow [0,1], x \mapsto \mu_{\tilde{a}}(x - b)$, leads to $\mathrm{COG}(\tilde{b}) = \mathrm{COG}(\tilde{a}) + b$. Hence, the COG moves the same width in the same direction.

– Rescaling the x-axis with a constant factor $c \in \mathbb{R} \setminus \{0\}$ should not change the relative position of the defuzzified value. For the COG this property is fulfilled, because the membership function $\mu_{\tilde{b}} : \mathbb{R} \longrightarrow [0,1], x \mapsto \mu_{\tilde{a}}\left(\frac{x}{c}\right)$, $c \in \mathbb{R} \setminus \{0\}$, yields $\mathrm{COG}(\tilde{b}) = c\,\mathrm{COG}(\tilde{a})$.

– On the other hand rescaling the degree of membership with a constant factor $c \in \mathbb{R}^+$ should be irrelevant to the defuzzified value. The membership function $\mu_{\tilde{b}} : \mathbb{R} \longrightarrow [0,1], x \mapsto c\mu_{\tilde{a}}(x)$, $c \in \mathbb{R}^+$, leads to $\mathrm{COG}(\tilde{b}) = \mathrm{COG}(\tilde{a})$. Hence, the property is fulfilled by the COG.

Many applications of fuzzy methodologies in the field of claims reserving in actuarial science make use of the following defuzzification method proposed by Campos Ibáñez and González Muñoz (1989) for fuzzy numbers.[12,13] As fuzzy numbers are convex fuzzy sets, the α-cuts are compact intervals. Hence, we will write for a fuzzy number $\tilde{a}$ the α-cuts as $A_\alpha = [\underline{A}_\alpha, \overline{A}_\alpha]$, $\underline{A}_\alpha, \overline{A}_\alpha \in \mathbb{R}$. The procedure is defined in the following way:

Definition 2.38

Let $\tilde{a}$ be a fuzzy number and $\beta \in [0,1]$. Let $A_\alpha = [\underline{A}_\alpha, \overline{A}_\alpha]$ denote the α-cuts of $\tilde{a}$ for all $\alpha \in [0,1]$. The expected value $E_\beta(\tilde{a})$ of a fuzzy number $\tilde{a}$ is defined as

$$E_\beta(\tilde{a}) := (1 - \beta) \int_0^1 \underline{A}_\alpha \, d\alpha + \beta \int_0^1 \overline{A}_\alpha \, d\alpha.$$

Remarks 2.39

a) The methods introduced in Chapters 5, 6 and 7 yield TFNs. The concept of an expected value of a TFN does not only offer a means to defuzzify the resulting values but the parameter provides an opportunity to assess the considered data.

[12] In fact, Campos Ibáñez and González Muñoz (1989) present a further approach to rank fuzzy numbers by an index introduced which can be also seen as a defuzzification technique.

[13] For the sake of comparison we choose the method by Campos Ibáñez and González Muñoz as recent publications concerning applications of fuzzy methods in claims reserving as e.g. Andrés Sánchez and Terceño Gómez (2003) and Andrés Sánchez (2006, 2007, 2012) use this defuzzification procedure.

With the help of the parameter β a decider can assess if he/she focuses more on $\overline{A}_\alpha$ or $\underline{A}_\alpha$. We refer to the parameter as "decision-maker risk parameter".

b) For a triangular fuzzy number $\tilde{a} = (a, l_a, r_a)$ the above definition simplifies to

$$\begin{aligned} E_\beta(\tilde{a}) &= (1-\beta)\left(a - \frac{l_a}{2}\right) + \beta\left(a + \frac{r_a}{2}\right) \\ &= a - \frac{1-\beta}{2} l_a + \frac{\beta}{2} r_a. \end{aligned}$$

This results directly when considering the functions to be integrated as depicted in Figure 2.12. For higher values of β more weight is put on the right spread.

c) Let $\tilde{b}$, $\tilde{c}$ be TFNs. $E_\beta(\tilde{a} \mid \cdot)$ denotes the expected value given a prior information. If the prior information is given by a set of TFNs $\{\tilde{b}, \tilde{c}, \ldots\}$, the TFNs shall be considered as crisp, i.e. $\{(b,0,0),(c,0,0),\ldots\}$.
It is assumed that there is no more uncertainty in the sense of fuzziness about the given information.

d) The expected value of a triangular fuzzy number $\tilde{a} = (a, l_a, r_a)$ possesses the following properties:

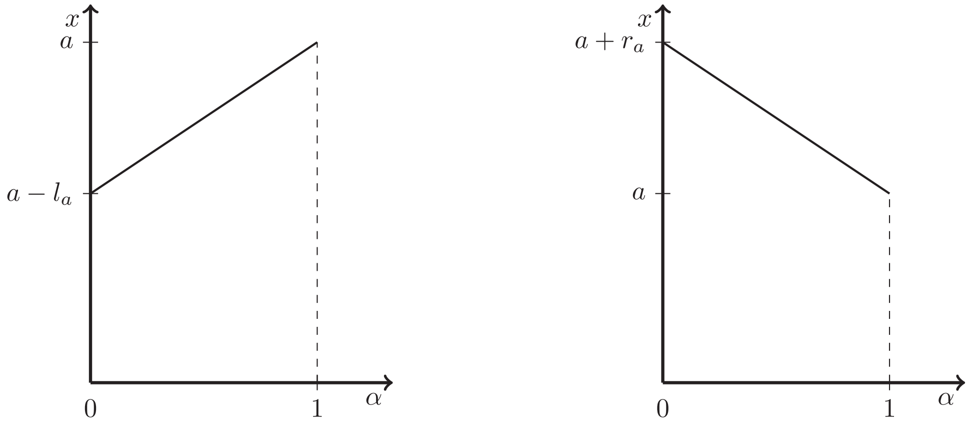
 - The expected value of a fuzzy number $\tilde{a}$ is always an element of the support of $\tilde{a}$ since

$$E_\beta(\tilde{a}) \in \left[a - \frac{l_a}{2}, a + \frac{r_a}{2}\right] \subset \operatorname{supp}(\tilde{a}).$$

 - For a crisp number $a \in \mathbb{R}$ the expected value yields the number itself, i.e. $E_\beta(a) = a$.

 - It is translation-invariant, i.e. when a constant $b \in \mathbb{R}$ is added to a triangular fuzzy number $\tilde{a}$ the expected value is shifted by this value in the direction of the x-axis, precisely

$$E_\beta(\tilde{a} \oplus (b,0,0)) = a + b - (1-\beta)\frac{l_a}{2} + \beta\frac{r_a}{2}.$$



(a) Function describing the lower bounds of the α-cuts of a triangular fuzzy number.

(b) Function describing the upper bounds of the α-cuts of a triangular fuzzy number.

Figure 2.12: Functions to be integrated when calculating the expected value of a triangular fuzzy number.

Weighted fuzzy arithmetic

As another defuzzification method the concept of weighted fuzzy arithmetic shall be mentioned which has been introduced by Chang (2001)[14]. It can be interpreted as a kind of mean value of a result of a fuzzy arithmetic operation (cf. Chang 2001, p. 226). The method is defined for normalized, asymmetric triangular membership functions. In case a normalized membership function is not at hand, it will be normalized.

Let $\tilde{a} = (a, l_a, r_a)$ and $\tilde{b} = (b, l_b, r_b)$ be asymmetric TFNs. Then, the α-cuts for a degree of membership $\alpha \in [0, 1]$ can be represented as[15]

$$^{\alpha}\tilde{a} := [^{\alpha}a_l, ^{\alpha}a_r] := [a - (1-\alpha)l_a, a + (1-\alpha)r_a]$$

and

$$^{\alpha}\tilde{b} := [^{\alpha}b_l, ^{\alpha}b_r] := [b - (1-\alpha)l_b, b + (1-\alpha)r_b].$$

With these notations weighted fuzzy addition and subtraction is defined in the following way (cf. Chang 2001, p. 226).

[14] Even though the name of the method suggests that it is a matter of arithmetical operations it is in fact a defuzzification procedure.

[15] In contrast to Definition 2.7 the notation of an α-cut is slightly changed for the sake of a better readability.

Definition 2.40 (Weighted fuzzy addition and subtraction)
Let $\tilde{a} = (a, l_a, r_a)$ and $\tilde{b} = (b, l_b, r_b)$ be asymmetric TFNs. The **weighted fuzzy sum** and the **weighted fuzzy difference** of $\tilde{a}$ and $\tilde{b}$ are defined as

$$\begin{aligned}\tilde{a} \oplus_{\mathrm{w}} \tilde{b} &:= \frac{\int_0^1 ({}^{\alpha}a_l + {}^{\alpha}b_l)\,\alpha\,\mathrm{d}\alpha + \int_0^1 ({}^{\alpha}a_r + {}^{\alpha}b_r)\,\alpha\,\mathrm{d}\alpha}{2\int_0^1 \alpha\,\mathrm{d}\alpha} \\ &= \int_{\alpha} ({}^{\alpha}a_l + {}^{\alpha}b_l)\,\alpha\,\mathrm{d}\alpha + \int_{\alpha} ({}^{\alpha}a_r + {}^{\alpha}b_r)\,\alpha\,\mathrm{d}\alpha\end{aligned}$$

and

$$\begin{aligned}\tilde{a} \ominus_{\mathrm{w}} \tilde{b} &:= \frac{\int_0^1 ({}^{\alpha}a_l + {}^{\alpha}b_l)\,\alpha\,\mathrm{d}\alpha - \int_0^1 ({}^{\alpha}a_r + {}^{\alpha}b_r)\,\alpha\,\mathrm{d}\alpha}{2\int_0^1 \alpha\,\mathrm{d}\alpha} \\ &= \int_{\alpha} ({}^{\alpha}a_l + {}^{\alpha}b_l)\,\alpha\,\mathrm{d}\alpha - \int_{\alpha} ({}^{\alpha}a_r + {}^{\alpha}b_r)\,\alpha\,\mathrm{d}\alpha.\end{aligned}$$

In Definition 2.40 the sum of the left borders and the sum of the right borders of the α-cuts are weighted with their grade of membership and integrated. The sum of both is put in relation to the integral of all values of α-cuts which is one because TFNs are normal.

Chang (2001) also defines weighted fuzzy operations for the multiplication and division of two TFNs (cf. Chang 2001, p. 227). As they are not used in the progress of this work we leave out their definitions.
Chang (2001) shows that Definition 2.40 can be simplified in the following way.

Proposition 2.41 (Weighted fuzzy addition)
Under the assumptions of Definition 2.40 the weighted fuzzy addition and subtraction can be represented as

$$\tilde{a} \oplus_w \tilde{b} = (a + b) + \tfrac{1}{6}(r_a + r_b - (l_a + l_b))$$

and

$$\tilde{a} \ominus_w \tilde{b} = (a + b) + \tfrac{1}{6}(r_a - r_b - (l_a - l_b)).$$

Proof. See Chang (2001, p. 227). □

Similar simplifications can also be derived for the other arithmetic operations (cf. Chang 2001, p. 227).

2.2.4 Measures of Fuzziness

The membership function is a measure which states the degree of membership for every single element of a fuzzy set. Nevertheless, there might be situations in which the user is not only interested in the vagueness of a single element but the total uncertainty of the fuzzy set. In order to further implement fuzzy sets in the analysis of situations dealing with vagueness a measure to compare the fuzziness of different sets is required. Moreover, an instrument to differentiate between varying severities of fuzziness is needed. While probability theory uses, among others, the variance as a measure to quantify the mean squared deviation to the mean and the mean squared error of prediction (MSEP, cf. Section 4.2) to specify the uncertainty of a prediction these means are not applicable here.

One of the first measures of fuzziness has been introduced by De Luca and Termini (cf. De Luca and Termini 1972).[16] A helpful overview of the most popular measures is given in Pal and Bezdek (1994).

In the following measures of fuzziness will be defined and preferable properties are detected. Subsequently, examples will be discussed. According to Pal and Bezdek (1994, p. 108) the definition of a measure of fuzziness is as follows.

Definition 2.42
A mapping

$$M: \quad F\mathscr{P}(X) \to [0, \infty)$$

is called **measure of fuzziness**.

Remarks 2.43

- The higher the value $M(\tilde{X})$ for a given fuzzy set $\tilde{X} \in F\mathscr{P}(X)$ the higher is the degree of fuzziness assigned to $\tilde{X}$.

- Some measures of fuzziness are motivated by the fact that the measure should quantify the deviation to a crisp set (cf. De Luca and Termini 1972). For a crisp set Y the measure of fuzziness yields zero, i.e. $M(Y) = 0$.

[16] Measures of fuzziness should not be mixed up with fuzzy measures.

Ebanks postulates a list of desirable properties a measure of fuzziness of a fuzzy set $\tilde{X} = \{(x, \mu_{\tilde{X}}(x)) \mid x \in X\}$ should fulfil (cf. Ebanks 1983, p. 25). Let $\tilde{Z} = \{(x, \mu_{\tilde{Z}}(x)) \mid x \in Z\}$ be also a fuzzy set. Then these properties[17] are:

(P1)	Sharpness	$M(\tilde{X}) = 0$ if and only if $\mu_{\tilde{X}}(x) \in \{0, 1\}$ for all $x \in X$, i.e. if $\tilde{X}$ is a crisp set.
(P2)	Maximality	If $\mu_{\tilde{X}}(x) = \frac{1}{2}$ for all $x \in X$, the function M attains its maximum.
(P3)	Resolution	When $\mu_{\tilde{Z}}(x) \leq \mu_{\tilde{X}}(x)$ if $\mu_{\tilde{X}}(x) \leq 0.5$ and $\mu_{\tilde{Z}}(x) \geq \mu_{\tilde{X}}(x)$ if $\mu_{\tilde{X}}(x) \geq 0.5$, i.e. $\tilde{Z}$ is a sharpened version of $\tilde{X}$, then $M(\tilde{X}) \geq M(\tilde{Z})$.
(P4)	Symmetry (to $\frac{1}{2}$)	$M(\tilde{X}) = M(\tilde{X}^C)$ holds true where $\mu_{\tilde{X}^C}(x) := 1 - \mu_{\tilde{X}}(x)$ for all $x \in X$.
(P5)	Valuation	The mapping M is called a valuation on $F\mathscr{P}(X)$ if $M(\tilde{X} \cup \tilde{Z}) + M(\tilde{X} \cap \tilde{Z}) = M(\tilde{X}) + M(\tilde{Z})$.

Remarks 2.44

- (P1) states that only for a crisp set the fuzziness is zero.
- The highest fuzziness is obtained for a fuzzy set in which all elements have a grade of membership of 0.5 as noted in (P2). In this case, the degree of membership of belonging or not-belonging to the set is identically high.
- If all grades of membership of elements are either closer to zero or one in comparison to another fuzzy set, the vagueness is reduced as there is less ambiguity about the membership to the set. Hence, $\tilde{Z}$ is called sharpened version in (P3) and it explains why sharpened versions should assign smaller fuzziness.
- Property (P4) means that a fuzzy set and its complement have the same grade of membership.
- According to property (P5) two fuzzy sets can exchange fuzziness (cf. Ebanks 1983, p. 26). However, the total fuzziness is given by the sum of the measures of fuzziness of the two sets.

[17] In fact, Ebanks (1983) states a further property (P6). Ebanks himself argues that property (P6) cannot be motivated intuitively and that it might be questionable. Therefore, we omit property (P6) here as it is not relevant for the further progress.

In the further progress of this work we need a measure of uncertainty for a fuzzy number, i.e. a measure of uncertainty for a fuzzy set with a piecewise continuous membership function over an interval of the real numbers $\mathbb{R}$. In order to do so we need to classify those measures which fulfill properties (P1)-(P5). In a first step this is done for the discrete finite setting as shown in Ebanks (1983).

Theorem 2.45 (Characterization of measures of fuzziness)
Let $X := \{x_1, \ldots, x_n\}$ and $M : F\mathscr{P}(X) \to [0, \infty)$ be a measure of fuzziness. Then M fulfills the properties (P1)-(P5) (cf. p. 44) if and only if M can be represented as $M(\tilde{A}) = \sum_{i=1}^{n} g(\mu_{\tilde{A}}(x_i))$, $\tilde{A} \in F\mathscr{P}(X)$, with a function $g : [0,1] \to \mathbb{R}_0^+$ for which holds:

i) $g(0) = g(1) = 0$ *and* $g(t) > 0 \quad \forall t \in (0,1)$

ii) $g(t) < g(0.5) \quad \forall t \in [0,1] \setminus \{0.5\}$

iii) *g is monotonically increasing on the interval $[0, 0.5)$ and monotonically decreasing on the interval $(0.5, 1]$*

iv) $g(t) = g(1-t) \quad \forall t \in [0,1]$

Proof. See Ebanks (1983, p. 28). □

With the help of Theorem 2.45 measures of fuzziness which fulfill properties (P1)-(P5) can be characterized. There exists a broad literature on those measures with various characteristics. An overview is given in Pal and Bezdek (1994) which also states which measure of fuzziness fulfils all (or parts) of the desirable properties (P1)-(P5) (cf. Pal and Bezdek 1994, Table 1, p. 111). Furthermore, Pal and Bezdek (1994) define two classes of measures of fuzziness whose members satisfy (P1)-(P5), the multiplicative and additive class. In the following we restrict ourselves to the multiplicative class because the measure of fuzziness which is used in the recent literature on applications of fuzzy methods in actuarial science, i.e. the expected value by Campos Ibáñez and González Muñoz (1989), belongs to that class.

Definition 2.46 (Multiplicative class)
Let $f : [0,1] \to \mathbb{R}_0^+$ be a differentiable, concave and increasing function, i.e. $f'(t) \geq 0$ and $f''(t) \leq 0$ for all $t \in [0,1]$, X be a finite set and $K \in \mathbb{R}^+$[18]. A measure of fuzziness M belongs to the **multiplicative class** if a representation

[18] Obviously, the constant K is not obligatory for the definition of the multiplicative class. Without any restrictions setting $K = 1$ also works. However, this definition is used in literature. By doing so different approaches of decision-makers can be incorporated.

$$M_{\text{mult}} = K \sum_{i=1}^{n} g(\mu_A(x_i)) \quad \text{for } A \in F\mathscr{P}(X)$$

exists where

$$g(t) := \hat{g}(t) - \min_{0 \le t \le 1} \hat{g}(t) \tag{2.10}$$

and

$$\hat{g}(t) := f(t)f(1-t). \tag{2.11}$$

Since measures of fuzziness with a representation as in Definition 2.46 satisfy the requirements of Theorem 2.45 they possess the desired properties (P1)-(P5).

Proposition 2.47
Measures of fuzziness which belong to the multiplicative class fulfill properties (P1)-(P5) (cf. p. 44).

Proof. See Pal and Bezdek (1994, p. 110). □

Pal and Bezdek (1994) extend the multiplicative class to fuzzy sets over intervals of real numbers.

Definition 2.48 (Multiplicative continuous class)
Let $X \subset \mathbb{R}$ and consider the same assumptions as in Definition 2.46. A measure of fuzziness belongs to the **multiplicative continuous class** if there exists a representation

$$M_{\text{mult,c}}(A) = K \int_X g(\mu_A(x))dx \quad \text{for } A \in F\mathscr{P}(X)$$

where the function g is given as in (2.10)-(2.11).

Proposition 2.49
Measures of fuzziness which belong to the multiplicative continuous class fulfill properties (P1)-(P4) (cf. p. 44).

Proof. See Pal and Bezdek (1994, p. 113). □

In the following we will define a measure of uncertainty for triangular fuzzy numbers since this is necessary in the progress of the work. We consider a triangular fuzzy number

$\tilde{a} = (a, l_a, r_a)$ with corresponding membership function as given in Definition 2.18. Let $f : [0,1] \to \mathbb{R}_0^+$, $f(t) = t\exp(1-t)$. Then f satisfies the assumption in Definition 2.46 since $f'(t) = (1-t)\exp(1-t) \geq 0$ for all $t \in [0,1)$ and $f''(t) = (t-2)\exp(1-t) \leq 0$ for all $t \in [0,1]$. Consequently, we have

$$\hat{g}(t) = f(t)f(1-t) = t\exp(1-t)(1-t)\exp(1-(1-t)) = t(1-t)e$$

and $\min_{0\leq t\leq 1} \hat{g}(t) = 0$. Hence, the functions g and $\hat{g}$ are equal, i.e. $g = \hat{g}$. Let $K \in \mathbb{R}^+$ then we yield a measure of fuzziness which fulfills properties (P1)-(P4) by:

$$\begin{aligned}
M_{\text{mult,c}}(\tilde{A}) &= K \int_X g(\mu_A(x))dx \\
&= K \int_x \mu_A(x)(1-\mu_A(x))e\,dx \\
&= K\Bigg(\int_{a-l_a}^{a} e\left(1+\frac{x-a}{l_a}\right)\left(1-1-\frac{x-a}{l_a}\right)dx \\
&\quad + \int_a^{a+r_a} e\left(1-\frac{x-a}{r_a}\right)\left(1-1+\frac{x-a}{r_a}\right)dx\Bigg) \\
&= K\left(-\frac{e}{l_a^2}\int_{a-l_a}^{a}(l_a+x-a)(x-a)dx + \frac{e}{r_a^2}\int_a^{a+r_a}(r_a-(x-a))(x-a)dx\right) \\
&= K\left(\frac{1}{6}e(l_a+r_a)\right) \\
&= \underbrace{K\frac{1}{3}e}_{=:K^*}\frac{1}{2}(l_a+r_a) \\
&= K^*\frac{1}{2}(l_a+r_a)
\end{aligned}$$

This motivates the following definition of a measure of uncertainty for triangular fuzzy numbers.[19] Furthermore, a definition for the case when previously known information is available is given. This will be of interest in the following Chapters 5, 6 and 7.

Definition 2.50 (Uncertainty of a TFN)

a) Let $\tilde{a} = (a, l_a, r_a)$ and $\tilde{b}$ be TFNs and $K \in \mathbb{R}^+$. The uncertainty $\text{Unc}_K(\tilde{a})$ of the TFN $\tilde{a}$ is given by:

[19] The definition is used in chapters 5, 6 and 7 and can be also found in Heberle and Thomas (2014).

$$\mathrm{Unc}_K(\tilde{a}) = \frac{1}{2}K(l_a + r_a)$$

b) $\mathrm{Unc}(\tilde{a} \mid \cdot)$ denotes the uncertainty of a TFN $\tilde{a}$ given a prior information. If the prior information is given by a set of TFNs $\{\tilde{b}, \tilde{c}, \ldots\}$, the TFNs shall be considered as crisp, i.e. $\{(b, 0, 0), (c, 0, 0), \ldots\}$.

Remarks 2.51

- The uncertainty of a TFN $\tilde{a} = (a, l_a, r_a)$ is independent of the mode a and depends only on the support of $\tilde{a}$.
- The uncertainty of a TFN $\tilde{a}$ can be interpreted as the (weighted) area between the x-axis and the membership function $\mu_{\tilde{a}}$. In fact, it is K times the area between the membership function and the x-axis. Thus, the larger the enclosed area, the larger the uncertainty Unc_K of a TFN $\tilde{a}$. In the same way a small area corresponds to a TFN with little uncertainty.
- In the context of the later chapters, the parameter K can be interpreted as follows: On the one hand larger values of the parameter K argue for little trust of the actuary whereas on the other hand smaller values of K stand for greater trust an actuary has in the data.
- The conditional uncertainty is motivated in analogy to the conditional expected value that there is no more uncertainty about known information in a fuzzy sense (cf. Definition 2.50b)).

2.3 Fuzzy Regression

In statistics and econometrics the classical linear regression model (see Appendix A) is used in a variety of applications. It sometimes might be applied even though some of the assumptions might be disturbed or the compliance might be difficult to verify. A classical linear regression model is of the form as presented in Definition A.1.

Classical linear regression methods are a collection of various estimation procedures and statistical tests in order to examine a relationship between a dependent variable and one (or more) independent variables. It is assumed that the relationship does not apply exact but might be a bit “noisy”. Therefore, an error term is added. In this way random errors are described but no systemic ones or imprecision in the data.[20] Moreover, the relationship between the dependent and the independent variables might be vague. These are situations which motivated the introduction of fuzzy regression techniques. In the literature there exist various different types of fuzzy regression models depending on the kind of fuzziness they like to address:

- fuzzy relationship in the data
- fuzzy input data
- fuzzy output data

Furthermore, the estimation methods divide the models in two parts. While in one part the estimation problem is transferred into a linear program, it is solved in others by a least squares approach.

An overview of fuzzy regression methods is given e.g. in Petry (1998), the miscellany of Kacprzyk and Fedrizzi (1992) and a discussion of the methods used in recent publications can be found in Shapiro (2006).

Considering the abundance of different fuzzy regression models only an assortment – chosen due to the progress of this work – is presented. The stated models differ in their estimation methods: the aim of the earlier published methods is to minimize the total fuzziness of the model. Their estimation problems can be reduced to a linear program (LP) and these methods are introduced in Section 2.3.1. A presentation of models using a L_2-distance measure to draw an analogy to classical regression can be found in Section 2.3.2.

2.3.1 Possibilistic Regression

A first model of fuzzy regression has been introduced by Tanaka et al. (1982) whose initial motivation is that the discrepancy between the observations and the estimated values “are assumed to depend on the indefiniteness of the system structure” (cf. Tanaka

[20] A description of possible sources of errors is presented in Schneeweiss and Mittag (1986, Introduction).

et al. 1982, p. 903).[21] They presume that the input data is crisp and the output data is fuzzy and that there exists a linear structure between that data.

Definition 2.52 (Fuzzy linear model)
Let $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $\tilde{a}_1, \ldots, \tilde{a}_n \in F\mathscr{P}(\mathbb{R})$ be triangular and symmetric, $\tilde{\mathbf{a}} := (\tilde{a}_1, \ldots, \tilde{a}_n)^T$ and $n \in \mathbb{N}$. The mapping

$$\begin{aligned} f: \quad & \mathbb{R}^n \longrightarrow F\mathscr{P}(\mathbb{R}) \\ & \tilde{y} := f(\mathbf{x}) := \tilde{a}_1 x_1 + \ldots + \tilde{a}_n x_n = \tilde{\mathbf{a}}^T \mathbf{x} \end{aligned}$$

is called **fuzzy linear model**.

Remarks 2.53

a) The fuzzy function f is defined and the membership function of $\tilde{y}$ in Definition 2.52 can be derived with the extension principle (cf. Section 2.1.3).

b) The fuzzy linear model can be extended by adding a fuzzy intercept $\tilde{a}_0 \in F\mathscr{P}(\mathbb{R})$ as it is done in later publications.

c) The fuzzy vector $\tilde{\mathbf{a}}$ shall be understood here component-wise, i.e. a vector containing fuzzy numbers as entries.

The fuzzy linear regression model for crisp input and output data has been introduced in Tanaka et al. (1980). In their 1982 model Tanaka et al. assume that the data can be represented by a fuzzy linear model and that the output data is fuzzy. Therefore, let $\mathbf{X} \in \mathbb{R}^{n\times m}$ a matrix of observations, $m, n \in \mathbb{N}$, $\mathbf{y} \in \mathbb{R}^n$. $\mathbf{x}_i := (x_{i1}, \ldots, x_{im})$ refers to the i-th row vector. Let $\tilde{y}_i := (y_i, e_i, e_i) \in F\mathscr{P}(\mathbb{R})$,$e_i \in \mathbb{R}$, $i = 1, \ldots, n$, be triangular and symmetric and $\tilde{\mathbf{y}} = (\tilde{y}_1, \ldots, \tilde{y}_n)$. The vagueness in the relationship between input and output data shall be expressed by fuzzy coefficients $\tilde{a}_1, \ldots \tilde{a}_m \in F\mathscr{P}(\mathbb{R})$, triangular and symmetric.
Given the above assumptions the authors assume that the data can be described by a fuzzy linear model

$$\tilde{y}_i = \tilde{a}_1 x_{i1} + \ \ldots + \tilde{a}_m x_{im} \qquad \text{for all } i = 1, \ldots, n.$$

[21]The model as well as the following model by Tanaka (1987) were initially proposed for fuzzy numbers of LR-type as coefficients. Since in this work all models are based on triangular fuzzy numbers the methods are presented only for these special cases.

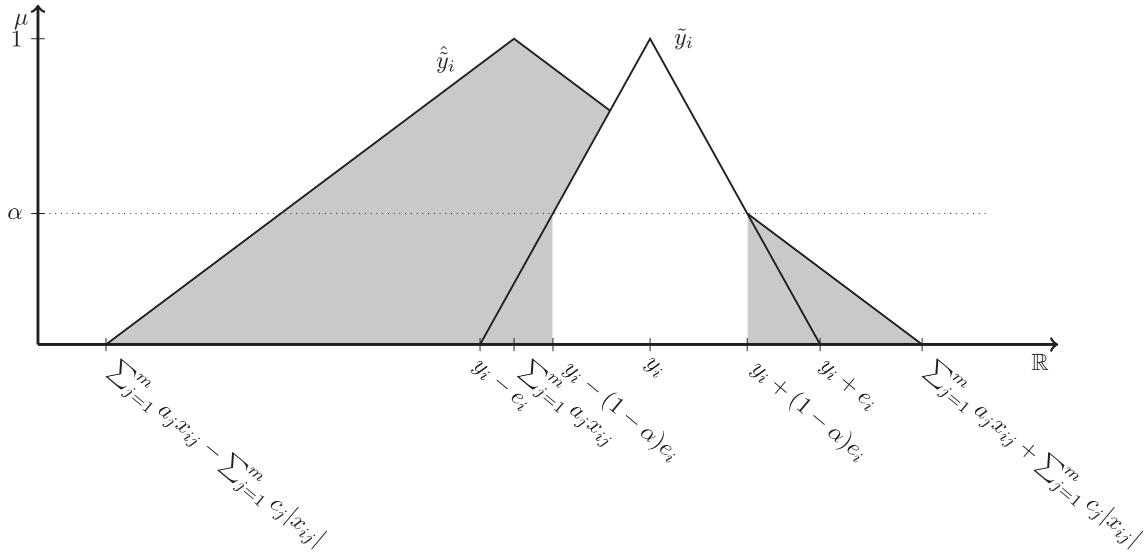


Figure 2.13: Simplified visualization of estimated fuzzy output $\hat{\tilde{y}}_i$ and observed fuzzy output $\tilde{y}_i$ for a fuzzy regression model.

We refer to $\hat{\tilde{y}}_i$, $i = 1, \ldots, n$, as the estimated output of the i-th input vector x_i and $\hat{\tilde{a}}_k$, $k = 1, \ldots, m$, as the estimations of the fuzzy coefficients, where $\hat{\tilde{a}}_k := (a_k, c_k, c_k)$, and $\mathbf{a} = (a_1, \ldots, a_m)^T, \mathbf{c} = (c_1, \ldots, c_m)^T$. Since the idea of this model is to minimize the total fuzziness by minimizing the sum of the spreads of the coefficients $\hat{\tilde{a}}_k$, $k = 1, \ldots, m$, the estimators of the fuzzy coefficients solve the following linear programming problem where $\alpha \in [0, 1]$.

$$\min_{c_1,\ldots,c_m} c_1 + \ldots + c_m \tag{2.12a}$$

s.t.

$$\mathbf{a}^T \mathbf{x}_i^T + (1-\alpha) \sum_{j=1}^m c_j |x_{ij}| \geq y_i + (1-\alpha) e_i, \quad i = 1, \ldots, n \tag{2.12b}$$

$$\mathbf{a}^T \mathbf{x}_i^T - (1-\alpha) \sum_{j=1}^m c_j |x_{ij}| \leq y_i - (1-\alpha) e_i, \quad i = 1, \ldots, n \tag{2.12c}$$

$$c_1, \ldots, c_m \geq 0 \tag{2.12d}$$

The first two constraints (2.12b)-(2.12c) assure that the given output is contained in the estimated output adjusted by the factor $(1 - \alpha)$, i.e. the α-cuts are considered. Tanaka et al. refer to the value α as H-value.[22] Choosing a reasonable value for α is of high interest and is extensively studied in Moskowitz and Kim (1993). Its interpretation

[22] In other publications it is also called h-certain factor.

is that the observed fuzzy output $\tilde{y}_i$ should lie within the support of the estimated fuzzy output $\hat{\tilde{y}}_i$ with a grade of membership of at least α. Values for the H-value range between zero and 0.9 (cf. Moskowitz and Kim 1993, p. 305).

An alternative model has then been proposed by Tanaka (1987) in which not the total spread of the coefficients as in (2.12) but the total width of the estimated output should be minimized. The linear problem here takes the form as shown in (2.13) (cf. Tanaka 1987, p. 367). The constraints (2.13b)-(2.13d) take the same form as in (2.12).

$$\min_{c_1,\ldots,c_m} \sum_{i=1}^{n}\sum_{j=1}^{m} c_j|x_{ij}| \tag{2.13a}$$

s.t.

$$\mathbf{a}^T\mathbf{x}_i^T + (1-\alpha)\sum_{j=1}^{m} c_j|x_{ij}| \geq y_i + (1-\alpha)e_i, \quad i = 1,\ldots,n \tag{2.13b}$$

$$\mathbf{a}^T\mathbf{x}_i^T - (1-\alpha)\sum_{j=1}^{m} c_j|x_{ij}| \leq y_i - (1-\alpha)e_i, \quad i = 1,\ldots,n \tag{2.13c}$$

$$c_1,\ldots,c_m \geq 0 \tag{2.13d}$$

Both models by Tanaka et al. (1982) and Tanaka (1987) describe a fuzzy relationship between dependent and independent variables without contemplating classical stochastic methods. The concept of the model of Tanaka (1987) is visualized in Figure 2.13. The observed fuzzy output $\tilde{y}_i$ should lie within the estimated fuzzy output $\hat{\tilde{y}}_i$ adjusted by the H-value. That is, the condition should hold for a grade of membership of at least α.

After publication of the first possibilistic regression models some criticism has arisen. Possibilistic regression models refer to these fuzzy regression models which are only using optimization methods as the two methods described before. Peters (1994) points out that the proposed methods are sensitive to outliers. Moreover, not all of the constraints in (2.12) and (2.13) are binding. In fact, most of them are obsolete and the bounds of the fuzzy regression interval are determined by the largest and smallest value, respectively (cf. Peters 1994, p. 94).[23] Wang and Tsaur (2000) criticize that there exists no convenient interpretation of the fuzzy regression interval.

[23]Ishibuchi and Nii (2001, p. 278) also argue that fuzzy linear models are only specified by a couple of data points. They also give an example of two data sets which are completely distributed differently but the four "outer" data points which are taken into account by the binding constraints are equal. Therefore the resulting fuzzy linear models are the same whereas the OLS estimates are different.

Since symmetric triangular fuzzy numbers are employed in those models they lack of variability. If the fuzziness in the relationship is more distinct to one side, i.e. either a stronger upward deviation or a downward one, this cannot be depicted by them.
Moreover, the fuzzy coefficients are determined by a linear program. Thus, there not necessarily exists a unique solution but all linear combinations of the basis solution also solve the problem.[24] In addition to that Bárdossy (1990) points out that the estimations are highly dependent on the "point-of-reference" of the input data. By ways of example he states the consideration of the difference of the data points and their mean value instead of the data points themselves. Beyond that Bárdossy extends the model to the use of different measures of vagueness.
Savic and Pedrycz (1991) highlight the fact that Tanaka et al. do not mention the issue of forecasting. Moreover, Diamond (1988) addresses the issue that the authors do not elaborate on the interrelation to classical linear models even though the title of the publication suggests something different (cf. Diamond 1988, pp. 141f.).[25]

In order to resolve some of the disadvantages of the early possibilistic regression models Ishibuchi and Nii (2001) proposed a model with asymmetric coefficients especially to eliminate those drawbacks caused by symmetric coefficients. It can be regarded as an extension of the model presented by Tanaka et al. (1982). In contrast to their model the modes of the coefficients are here estimated by OLS.[26] The coefficients are supposed to be asymmetric triangular fuzzy numbers, i.e. $\tilde{a}_i := (a_i, l_{a_i}, r_{a_i}) \in F\mathscr{P}(\mathbb{R})$, triangular, $i = 0, \ldots, m$. Then $\tilde{\mathbf{a}} = (\tilde{a}_0, \ldots, \tilde{a}_m)^T \in F\mathscr{P}(\mathbb{R}^{m+1})$ denotes the fuzzy coefficient vector. Let

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots & \\ 1 & x_{n1} & \cdots & x_{nm} \end{pmatrix} \in \mathbb{R}^{n\times(m+1)}$$

a matrix of observations and $\tilde{\mathbf{y}} \in F\mathscr{P}(\mathbb{R}^n)$ a fuzzy output vector which can be either fuzzy, i.e. represented by an interval, or crisp. Ishibuchi and Nii consider the following fuzzy linear model:

[24] Petry (1998, pp. 84ff.) gives examples in which the variety of solutions is shown.
[25] A comparison of the two presented methods with OLS has been conducted by Redden and Woodall (1994) and Chang and Ayyub (2001). They also point out the drawbacks of the 1980 and 1982 methods of Tanaka et al.
[26] The same approach of determining the center by OLS is pursued by Savic and Pedrycz (1991).

$$
\begin{aligned}
\tilde{y}_i &= (y_i, l_{y_i}, r_{y_i}) \\
&= \tilde{a}_0 \oplus \tilde{a}_1 x_{i1} \oplus \ldots \oplus \tilde{a}_m x_{im} \\
&= (a_0, l_{a_0}, r_{a_0}) \oplus (a_1, l_{a_1}, r_{a_1}) x_{i1} \oplus \ldots \oplus (a_m, l_{a_m}, r_{a_m}) x_{im} \\
&= (a_0, l_{a_0}, r_{a_0}) \oplus \left(a_1 x_{i1}, \begin{cases} l_{a_1} x_{i1} & \text{if } x_{i1} \geq 0 \\ r_{a_1} |x_{i1}| & \text{if } x_{i1} < 0 \end{cases}, \begin{cases} r_{a_1} x_{i1} & \text{if } x_{i1} \geq 0 \\ l_{a_1} |x_{i1}| & \text{if } x_{i1} < 0 \end{cases} \right) + \ldots \\
&\quad + \left(a_m x_{im}, \begin{cases} l_{a_m} x_{im} & \text{if } x_{im} \geq 0 \\ r_{a_m} |x_{im}| & \text{if } x_{im} < 0 \end{cases}, \begin{cases} r_{a_m} x_{im} & \text{if } x_{im} \geq 0 \\ l_{a_m} |x_{im}| & \text{if } x_{im} < 0 \end{cases} \right), \qquad i = 1, \ldots, n
\end{aligned}
$$

Thus, we yield for the mode y_i and the left and right spread l_{y_i} and r_{y_i}, respectively:

$$
\begin{aligned}
y_i &= a_0 + \sum_{j=1}^{m} a_j x_{ij} \\
l_{y_i} &= l_{a_0} + \sum_{\substack{j=1 \\ x_{ij} \geq 0}}^{m} l_{a_j} x_{ij} + \sum_{\substack{j=1 \\ x_{ij} < 0}}^{m} r_{a_j} |x_{ij}| \\
r_{y_i} &= r_{a_0} + \sum_{\substack{j=1 \\ x_{ij} \geq 0}}^{m} r_{a_j} x_{ij} + \sum_{\substack{j=1 \\ x_{ij} < 0}}^{m} l_{a_j} |x_{ij}|
\end{aligned}
$$

We refer to the estimated output as $\hat{\tilde{y}}_i$ and denote the mode by $\hat{y}_i$ and the estimated left and right spreads as $l_{\hat{y}_i}$ and $r_{\hat{y}_i}$, respectively. An α-cut (cf. Definition 2.7), $\alpha \in [0, 1]$, of the estimated output is determined by

$$
\left(\hat{\tilde{y}}_i\right)_\alpha = [\hat{y}_i - (1 - \alpha) l_{\hat{y}_i}, \hat{y}_i + (1 - \alpha) r_{\hat{y}_i}].
$$

Ishibuchi and Nii propose a two-step procedure in their estimation method. First, the modes of the coefficients $a_0, \ldots, a_m$ are estimated with OLS. They are denoted in the following by $\hat{a}_0, \ldots, \hat{a}_m$. Then, we yield the left and right spreads of the estimated coefficients by solving the following linear model (2.14) which aims to minimize the total spread of the estimated output.

$$
\min_{\hat{l}_{a_0}, \ldots, \hat{l}_{a_m}, \hat{r}_{a_0}, \ldots, \hat{r}_{a_m}} \left(\sum_{j=0}^{m} \sum_{i=1}^{n} \hat{l}_{a_j} |x_{ij}| + \hat{r}_{a_j} |x_{ij}| \right) \tag{2.14a}
$$

s.t.

$$\hat{a}_0 + \sum_{j=1}^{m} \hat{a}_j x_{ij} - \left(\hat{l}_{a_0} + \sum_{\substack{j=1 \\ x_{ij} \geq 0}}^{m} \hat{l}_{a_j} |x_{ij}| + \sum_{\substack{j=1 \\ x_{ij} < 0}}^{m} \hat{r}_{a_j} |x_{ij}| \right) (1-\alpha) \leq y_i, \quad i = 1, \ldots, n \tag{2.14b}$$

$$\hat{a}_0 + \sum_{j=1}^{m} \hat{a}_j x_{ij} + \left(\hat{r}_{a_0} + \sum_{\substack{j=1 \\ x_{ij} \geq 0}}^{m} \hat{r}_{a_j} |x_{ij}| + \sum_{\substack{j=1 \\ x_{ij} < 0}}^{m} \hat{l}_{a_j} |x_{ij}| \right) (1-\alpha) \geq y_i, \quad i = 1, \ldots, n \tag{2.14c}$$

$$\hat{l}_{a_0}, \ldots, \hat{l}_{a_m}, \hat{r}_{a_0}, \ldots, \hat{r}_{a_m} \geq 0 \tag{2.14d}$$

The constraints (2.14b)-(2.14c) assure that the observed output is contained in the estimated one. Hereby Ishibuchi and Nii (2001) assume that the observed output data y_i is crisp.[27] Constraint (2.14d) ensures that the left and right spreads are non-negative.

2.3.2 Fuzzy Least Squares Regression

A second large part of fuzzy regression models can be classified as fuzzy least squares regression models. Following the idea of ordinary least squares in classical statistics the method is based on a distance measure between the fuzzy or non-fuzzy input variables and the fuzzy output. The aim is to derive parameters by minimizing the quadratic distance.

One of the most popular representatives of this class is the model by Diamond and Kloeden (1994) which has also been regarded in Diamond (1992). Further considerations can be found among others in Diamond and Tanaka (1998) and Petry (1998). The presentation of the model of Diamond and Kloeden (1994) in this section is based on the latter. This technique comprises both non-fuzzy and fuzzy input data. Merely, the fuzzy data needs to be in the form of triangular fuzzy numbers. As a difference to other parts of this work we will denote TFNs $\tilde{a} = (a, l_a, r_a)$ in the form $\tilde{a} =: [a - l_a, a + r_a] =: [a^l, a^r]$

[27] Their model can be extended to the case with observed fuzzy output data by replacing y_i by the terms $y_i - (1-\alpha)l_{y_i}$ and $y_i + (1-\alpha)r_{y_i}$, respectively, in analogy to their symmetric case (cf. Ishibuchi and Nii 2001, pp. 276f.).

in this section, i.e. not the width of the spreads are being denoted but the lower and upper bound of the support.

The goal is to minimize the distance between the observed values – either fuzzy or non-fuzzy– and fuzzy values which are given by a parametrical model (Diamond and Tanaka 1998, p. 353). This metric on the space of TFNs is defined as

$$\begin{aligned} D_2(\tilde{a}, \tilde{b}) &= D_2\left([a^l, a^r], [b^l, b^r]\right) \\ &:= \sqrt{(a^l - b^l)^2 + (a^r - b^r)^2} \end{aligned}$$

in Diamond and Kloeden (1994, p. 115) where $\tilde{a}$ and $\tilde{b}$ are TFNs.

Fuzzy Least Squares approach by Diamond and Kloeden (1994)
The following fuzzy linear regression model will be considered: Let $\tilde{a}$, $\tilde{b}$ be TFNs, $\tilde{y}_i = (y_i, l_{y_i}, r_{y_i})$, $i = 1, \ldots, n$, be fuzzy output and x_i, $i = 1, \ldots, n$, non-fuzzy input data.[28]

$$\begin{aligned} & & \tilde{y}_i &= \tilde{a} \oplus \tilde{b} x_i \\ \Leftrightarrow & & [y_i^l, y_i^r] &= [a^l, a^r] \oplus [b^l, b^r] x_i \end{aligned} \tag{2.15}$$

That is, the notation $[y_i^l, y_i^r]$ as well describes the upper and lower bound of the support. For the model as stated in Equation (2.15) it is required that the slope parameter $\tilde{b}$ is either positive or negative, i.e. the support of $\tilde{b}$, $\mathrm{supp}(\tilde{b})$, does not contain zero.[29] There are no restrictions for the parameter $\tilde{a}$.

Following the idea of OLS in classical statistics estimators for the fuzzy coefficients $\tilde{a}$ and $\tilde{b}$ of the model as stated in Equation (2.15) are derived by minimizing the quadratic distances between the fuzzy regression line and the observed data, i.e.

$$\min_{\substack{\tilde{a}, \tilde{b} \\ 0 \notin \mathrm{supp}(\tilde{b})}} \sum_{i=1}^{n} D_2\left([a^l, a^r] \oplus [b^l, b^r] x_i, [y_i^l, y_i^r]\right)^2 \tag{2.16}$$

[28] The input data can be also given as fuzzy data. For an easier readability we will stick to the crisp case here but keep in mind that it can be defined for fuzzy input data as well.
[29] This fact is in detail discussed in Diamond and Kloeden (1994, pp. 117f.) and Petry (1998, pp. 94f.). It is required for the uniqueness of the solution of the minimization problem.

In Diamond (1992) and Diamond and Kloeden (1994) the existence and uniqueness of the solution of the minimization problem as stated in Equation (2.16) is shown (see Diamond and Kloeden (1994, pp. 117f.), a plausibility check is also given in Petry (1998, pp. 94f.)).[30]

Early publications in the field of fuzzy least squares regression can be found in Celmiņš (1987a) and Celmiņš (1987b). The presented model by Diamond and Kloeden (1994) then picks up the idea of OLS in classical statistics and transfers it to fuzzy regression. The thought is further pursued in Körner and Näther (1998) who investigated fuzzy random variables (for a short introduction see Section 2.4) in the context of fuzzy linear regression. A least squares approach with fuzzy random variables is also considered in Wünsche and Näther (2002).

A further fuzzy regression model making use of a least squares approach which came up during the course of years is given by Chang (2001). Chang (2001) introduces a hybrid fuzzy least squares regression method making use of weighted fuzzy arithmetic (cf. p. 41). The fuzzy regression method is defined both for a bivariate and a multivariate model. As we are only dealing with the bivariate model in Section 4.6.5 we restrict ourselves to that case.

Hybrid Fuzzy Least Squares Regression Model by Chang (2001)
A bivariate fuzzy regression model of the following manner is considered

$$\tilde{y}_i = \tilde{a}_0 \oplus \tilde{a}_1 x_i, \qquad i = 1, \dots n,$$

where the fuzzy coefficients are TFNs, i.e. $\tilde{a}_0 := (a_0, l_{a_0}, r_{a_0})$ and $\tilde{a}_1 := (a_1, l_{a_1}, r_{a_1})$. $\tilde{y}_i$ denotes the fuzzy output and x_i, $i = 1, \dots, n$, the crisp input variable. Consequently, the fuzzy dependent variable can be displayed in the following form:

$$\tilde{y}_i = (a_0, l_{a_0}, r_{a_0}) \oplus (a_1, l_{a_1}, r_{a_1}) x_i = (a_0 + a_1 x_i, l_{a_0} + l_{a_1} x_i, r_{a_0} + r_{a_1} x_i)$$

[30] Diamond and Kloeden (1994) goes into detail with *cohesive* data which corresponds to the requirements for the uniqueness of the regression model's solution (cf. Diamond and Kloeden 1994, p. 117). This view is also described in detail in Petry (1998). Vividly speaking, the constraint for the uniqueness of the solution means that the lines describing the left and right spread of the fuzzy regression tube do not intersect. In that case the fuzzy numbers would not be defined (cf. Petry 1998, p. 95). If the data is not *cohesive* a solution does also exist and is unique.

Then, the lower and upper borders of the α-cuts, $\alpha \in [0,1]$, of the estimated values $\hat{y}_i$ which are referred to as ${}^{\alpha}\hat{y}_{i,l}$ and ${}^{\alpha}\hat{y}_{i,r}$ are given by

$$ {}^{\alpha}\hat{y}_{i,l} = a_0 + a_1 x_i - (1-\alpha)(l_{a_0} + l_{a_1} x_i) $$

and

$$ {}^{\alpha}\hat{y}_{i,r} = a_0 + a_1 x_i + (1-\alpha)(r_{a_0} + r_{a_1} x_i). $$

In the following we will denote by $\tilde{y}_i = (y_i, l_{y_i}, r_{y_i})$, $i = 1, \ldots, n$, the observed values of the dependent variable for which the lower and upper borders of the corresponding α-cuts ${}^{\alpha}\tilde{y}_{i,l}$ and ${}^{\alpha}\tilde{y}_{i,r}$ are given by

$$ {}^{\alpha}\tilde{y}_{i,l} = y_i - (1-\alpha) l_{y_i} $$

and

$$ {}^{\alpha}\tilde{y}_{i,r} = y_i + (1-\alpha) r_{y_i}. $$

The idea of the hybrid regression model by Chang (2001) is to minimize the sum of squared residual errors in analogy to the procedure in OLS. Hence, the objective is to minimize the subsequent term:

$$ \sum_{i=1}^{n} (\hat{y}_i \ominus_{\mathrm{w}} \tilde{y}_i)^2 = \sum_{i=1}^{n} \frac{\int_0^1 ({}^{\alpha}\hat{y}_{i,l} -{}^{\alpha} \tilde{y}_{i,l})\, \alpha \,\mathrm{d}\alpha + \int_0^1 ({}^{\alpha}\hat{y}_{i,r} -{}^{\alpha} \tilde{y}_{i,r})\, \alpha \,\mathrm{d}\alpha}{2\int_0^1 \alpha \,\mathrm{d}\alpha}, $$

where the operation $\ominus_{\mathrm{w}}$ denotes the subtraction according to the definition of weighted fuzzy arithmetic (cf. Definition 2.40). Chang shows that the sum of squared residual errors can be represented as (cf. Chang 2001, pp. 228ff.):

$$ \begin{aligned} \sum_{i=1}^{n} (\hat{y}_i \ominus_{\mathrm{w}} \tilde{y}_i)^2 = \sum_{i=1}^{n} \Big[& (a_0 + a_1 x_i - y_i)^2 \\ & + \tfrac{1}{3}(a_0 + a_1 x_i - y_i)\Big((r_{a_0} + r_{a_1} x_i - r_{y_i}) - (l_{a_0} + l_{a_1} x_i - l_{y_i})\Big) \\ & + \tfrac{1}{12}\Big((l_{a_0} + l_{a_1} x_i - l_{y_i})^2 + (r_{a_0} + r_{a_1} x_i - r_{y_i})^2\Big)\Big] \\ =: {} & F(a_0, a_1, l_{a_0}, l_{a_1}, r_{a_0}, r_{a_1}) \end{aligned} $$

In order to determine the set of coefficients $\{a_0, a_1, l_{a_0}, l_{a_1} r_{a_0}, r_{a_1}\}$ which minimize the function F the partial derivatives of F need to be derived and subsequently set to zero. Chang demonstrates that instead of solving the resulting linear system of equations with six equations and six unknowns the following three linear systems of equations

(2.17)-(2.19) with two equations and two unknowns each can be solved (cf. Chang 2001, pp. 230 ff.).

$$\left| \begin{array}{rcl} na_0 & + \; a_1 \sum_{i=1}^n x_i = & \sum_{i=1}^n y_i \\ a_0 \sum_{i=1}^n x_i & + \; a_1 \sum_{i=1}^n x_i^2 = & \sum_{i=1}^n x_i y_i \end{array} \right| \quad \text{linear system of equations for } a_0 \text{ and } a_1 \tag{2.17}$$

$$\left| \begin{array}{rcl} nl_{a_0} & + \; l_{a_1} \sum_{i=1}^n x_i = & \sum_{i=1}^n l_{y_i} \\ l_{a_0} \sum_{i=1}^n x_i & + \; l_{a_1} \sum_{i=1}^n x_i^2 = & \sum_{i=1}^n x_i l_{y_i} \end{array} \right| \quad \text{linear system of equations for } l_{a_0} \text{ and } l_{a_1} \tag{2.18}$$

$$\left| \begin{array}{rcl} nr_{a_1} & + \; r_{a_1} \sum_{i=1}^n x_i = & \sum_{i=1}^n r_{y_i} \\ r_{a_0} \sum_{i=1}^n x_i & + \; r_{a_1} \sum_{i=1}^n x_i^2 = & \sum_{i=1}^n x_i r_{y_i} \end{array} \right| \quad \text{linear system of equations for } r_{a_0} \text{ and } r_{a_1} \tag{2.19}$$

In case of symmetric fuzzy regression coefficients and symmetric fuzzy output variables only the first two systems of equations (2.17) and (2.18) need to be solved since the third one is obsolete.

2.4 Fuzzy Random Variables

Univariate or multivariate random variables in its classical sense are mappings from a probability space $(\Omega, \mathcal{A}, P)$ into $\mathbb{R}$ or $\mathbb{R}^k$, respectively, which are measurable. Again, in this approach solely stochastic randomness is taken into account (see Section 2.1.4). The concept of fuzzy random variables allows for modeling both stochastic randomness and fuzziness, e.g. emerging from linguistic uncertainty.

The idea came up in the works of Kwakernaak (1978, 1979) and Puri and Ralescu (1986). Kruse and Meyer (1987) later on specified the ideas of Kwakernaak (1978, 1979). It has also been considered by Diamond and Kloeden (1994) with a different approach. An attempt of a unification of those different views and definitions has been made by Krätschmer (2001). An overview – with an emphasis on insurance – can be found in Shapiro (2009) and a presentation of the development as well as further interpretations are given in Gil et al. (2006). The presentation in this section is based on Shapiro (2009).

As an example of a fuzzy random variable we will briefly state the definition of Kwakernaak's view of a fuzzy random variable who sees them as "[...] random variables whose values are not real but fuzzy numbers [...]" (cf. Kwakernaak 1978, p. 1). In this case they are modeled as a fuzzy conception of a crisp random variable which is unobservable.

Definition 2.54 (Fuzzy Random Variable (Kwakernaak))
Let $(\Omega, \mathcal{A}, P)$ be a probability space and $F\mathscr{P}(\mathbb{R})$ the set of all fuzzy numbers. That is, $F\mathscr{P}(\mathbb{R})$ comprises all normal and convex fuzzy sets whose α-cuts are compact, $\alpha \in [0,1]$. Specifically, this is the class of mappings $\zeta : \mathbb{R} \longrightarrow [0,1]$ for which we define

$$\zeta_\alpha := \begin{cases} \{x \in \mathbb{R} \mid \zeta(x) \geq \alpha\} & \text{if } \alpha \in (0,1] \\ \overline{\operatorname{supp}(\zeta)} & \text{if } \alpha = 0 \end{cases},$$

where $\overline{\operatorname{supp}(\zeta)}$ denotes the closure of the support of ζ. A fuzzy random variable ξ is a mapping $\xi : (\Omega, \mathcal{A}, P) \longrightarrow F\mathscr{P}(\mathbb{R})$ such that for all $\omega \in \Omega$ and $\alpha \in [0,1]$ the mappings

$$\inf \xi_\alpha : (\Omega, \mathcal{A}, P) \longrightarrow \mathbb{R}, \qquad \inf \xi_\alpha(\omega) := \inf(\xi(\omega))_\alpha$$

and

$$\sup \xi_\alpha : (\Omega, \mathcal{A}, P) \longrightarrow \mathbb{R}, \qquad \sup \xi_\alpha(\omega) := \sup(\xi(\omega))_\alpha$$

are real-valued random variables, i.e. measurable functions.
The above conditions on the boundaries of the fuzzy random variable can also be written as

$$\xi_\alpha(\omega) = [\inf(\xi(\omega))_\alpha, \sup(\xi(\omega))_\alpha].$$

Remark 2.55
The requirement for measurable functions in Definition 2.54 ensures that the theory of well examined random variables can be applied to investigate fuzzy random variables. The mapping ξ goes into the space of all fuzzy numbers which are normal by definition. This assures that the image is not empty. According to Petry (1998, p. 30) the definition depends on the idea that there exists a “real” random variable describing an experiment which cannot be perceived precisely. The grades of membership are then interpreted as a grade of acceptance that the value is the “true” realization of the random variable.

3 Applications of Fuzzy Theory in Insurance

It is an essential characteristic of the insurance industry that it is exposed to many sources of randomness and imprecision. For example this may involve the fields of pricing, underwriting, risk classification and reserving. A key goal has always been to quantify the related uncertainty in order to e.g. price an insurance product adequately or to set up an appropriate reserve.

While classical stochastic methods have been used for decades it took more than 15 years since the appearance of the first article in fuzzy set theory by Zadeh (1965) that the methodology of fuzzy sets has found its way into insurance related fields. To our best knowledge, the first article was published in 1982 by Wit. Derrig and Ostaszewski (1999) try to find answers why it took so long for researchers to find suitable implementations. Their answer is mainly valid for the USA as the authors mention that the situation in e.g. Europe is slightly different (cf. Derrig and Ostaszewski 1999, p. 534). Strict regulations have led to the situation that actuarial calculations were conducted following the guidelines (cf. Derrig and Ostaszewski 1999, p. 534) and other techniques have not necessarily been considered in those times. Since then numerous articles in various areas of application in insurance have been published. The use of fuzzy set theory ranges from different product lines as property-liability to health and life insurances, from underwriting and reinsurance decisions to financial analyses, from the mere fuzzy logic to fuzzy clustering methods and fuzzy regression.

A survey can among others be found in Shapiro (2004), Shapiro (2007) or Yakoubov and Haberman (1998). An overview is also given in the introductions of the articles by

Andrés Sánchez and Terceño Gómez (2003) and Andrés Sánchez (2006) as well as in Shapiro and Jain (2003, Chapter 1).

In the following four sections applications of fuzzy techniques in the fields of underwriting, pricing, risk classification and claims reserving are presented. We have exemplarily chosen underwriting, pricing and risk classification as these are adjacent fields to claims reserving. Hereby we focus on those publications considering non-life insurance applications. However, the publications dealing with life insurances applications which have been important in the development are also mentioned. The chosen fields of application are not extensive. Additionally, fuzzy methods are used in e.g. fraud detection and insurer surveillance (cf. Berry-Stölzle et al. 2010). Moreover, asset allocation can be of great interest in an insurance context. Also the overview by Shapiro (2004) refers mainly to finance literature when considering asset allocation so that we will exclude it in this dissertation. The summary in the following sections does not aim to be comprehensive but rather gives an impression of how different applications of fuzzy methods in insurance are.

3.1 Underwriting

Underwriting in an insurance company deals with the selection of risks which will be signed by the insurer as well as the conditions and shares of signed risks. The superior goal is to have a profitable portfolio of risks.

The – to our knowledge – first publication of an application of fuzzy methods at all considers an application in underwriting and was published in the early 1980's (cf. Wit 1982). The author's motivation is that the underwriting process is subjective when assessing risks. In case a rating system is at hand, employees in the underwriting department might use it; but maybe according to the underwriter's intuition. For an underwriter human factors play an important role such that they make use of their intuition and also utilize incomplete information (cf. Wit 1982, p. 278). Wit assigned membership functions by defining membership grades for linguistic terms for several areas. Then, the combination of the assigned fuzzy values is considered in order to derive the underwriting decision. This particular idea of using fuzzy methods in underwriting is further pursued by Young (1993).

Young (1996) applies a model to group health insurance. As the investigation is so general the findings can be applied to any kind of LoB and, thus, it is also incorporated in the overview. The aim of the publication is to show a way to consider vague or linguistic objectives as well. Fuzzy constraints are used to adjust insurance premiums. With the help of this means premiums can be altered even if experience data is not available. For doing so, fuzzy rules are applied. These rules can be operations that correspond to the linguistic connectors "and" and "or" as well as the modifier "not". Young sees a major advantage of her work in the fact that an investigation can start off with verbal rules (cf. Young 1996, p. 482). Subsequently, fuzzy sets describing the hypotheses are constructed (cf. Young 1996, p. 469). Then, solutions are derived and an option of fine-tuning of the fuzzy logic models is also discussed.

The publication of Lemaire (1990) aims to give an adaptable definition of a preferred policy holder in a life insurance context. The author takes into account blood pressure, cholesterol level, weight and the status whether the policy holder is a smoker or non-smoker. For each criteria a membership function is defined (cf. Lemaire 1990, p. 38). For example for weight, the membership function describes a weight which is between 85 % and 130 % of the recommended one. A preferred policy holder is then the intersection of the fuzzy sets which are described by the membership functions. Moreover, Lemaire extends the literature by broadening the definition of the intersection.

The first application of fuzzy methods in underwriting outside academia is to our best knowledge by a Canadian team who used among others fuzzy techniques to develop an automated underwriter (mentioned e.g. in Erbach and Seah 1993). Horgby et al. (1997) derive an application to medical underwriting in life insurance in the case of applicants suffering from diabetes mellitus. The authors utilize a fuzzy inference system. One of the reasons for doing so is that it is practicable for physicians to assess a symptom being mild or severe but it might be not so useful to quantify the severity.

The work of Bonissone (2004) deals with the process of underwriting insurance applications and considers the problem of risk classification. He provides a design methodology for a fuzzy knowledge-based classifier. Bonissone (2003) as well describes applications for the underwriting process. The author considers the whole process with particular regard to the maintenance of a model. The work of Bonissone et al. (2005) extends previous publications by providing a model for the underwriting process in a dynamic environment which shows a better performance than static ones. It is presented for an example of a fleet selection problem. The process of underwriting in the context of workers compensation insurance for construction workers is addressed in Imriyas et al. (2006, 2007) and Imriyas (2009).

3.2 Pricing

Pricing is the process of assigning an adequate price for a given risk. The classical approach is to calculate the claims expectancy. Nowadays, there is a wide range of methods – even including ideas from behavioral economics (cf. Schmidt-Gallas 2014).

As elaborated on before, Young (1996) showed how fuzzy logic could be used in insurance questions. Moreover, the designed model can be of use in pricing matters. The extension of the article is shown in Young (1997). The publication makes use of two sources of information: data on claims (experience data) and additional information which can be financial or marketing data as well as documents showing the actuary's or company's philosophy. Starting off from linguistic rules Young develops a fuzzy inference system to adjust premiums (cf. Young 1997, p. 735) – including a step to fine-tune the model. The earlier model only takes into account constraints. Her model is then applied to a workers compensation data set.

Cummins and Derrig (1997) use fuzzy set theory in the context of financial pricing of property-liability insurance contracts. Especially, the authors apply fuzzy methods to the Myers-Cohn model, a discrete-time discounted-cash-flow model (cf. Cummins and Derrig 1997, p. 21). According to Cummins and Derrig fuzzy set theory allows for an approach to comprise vague, fuzzy or incomplete information (cf. Cummins and Derrig 1997, p. 22). In pricing this might be the case for information on cash flows, future economic conditions, risk premiums or other aspects concerning the pricing decision. The authors derive numerical examples and come to the conclusion that considering fuzzy premiums offers other information than taking into account crisp ones. In fact, fuzzy set theory allows to incorporate vagueness and uncertainty "bottom up" instead of "top down" (cf. Cummins and Derrig 1997, p. 37).

Carretero and Viejo (2000) introduce a fuzzy Bonus Malus system for automobile insurance. Their study is based on a data set of 48,666 policy holders of an insurance company in Spain. Bonus Malus systems recompense drivers for claims-free years and add an extra premiums for years with at least one claim. The investigation makes use of fuzzy methods in a decision-making process. The objective function is to maximize earnings from premiums while the principle is that every insured pays a premium proportional to his/her claim frequency. Moreover, the system is supposed to be financially balanced in a way that premiums can cover expected claims and there is no big cluster in a highly discounted class. In comparison to the crisp case the fuzzy Bonus Malus system introduces qualitative constraints, i.e. constraints with linguistic variables and, therefore, yields more flexibility.

A recent publication of Luukka and Collan (2015) deals with the pricing of large projects' insurances, especially large industrial investments. In order to model the investment's profitability the authors employ possibility theory which is a subfield of fuzzy set theory. Based on these considerations they develop an approach for the pricing of the project's insurance. The accessible information for the planning and analysis phase is often vague (cf. Luukka and Collan 2015, p. 23). That is the reason why again possibility theory is utilized for pricing giga-investments' insurances by considering the pay-off distributions of the projects (cf. Luukka and Collan 2015, p. 28). Furthermore, the effects of risk aversion in the setting of giga-investments is investigated in Collan et al. (2016).

Yao and Qin (2015) present an application to risk processes with uncertain factors which are estimated by subjective factors and assessment. Lai (2006) fits fuzzy triangular numbers to the ICAPM model which is a variation of the well-known CAPM model for property-liability insurers.

The publications mentioned before all focus on the supply side of the relationship between an insurer and the insured. In the following articles taking into consideration the demand side are presented. Nonetheless, considering the demand side also has an impact for those who set the price. Liu et al. (2015) model the potential loss of an insured in an optimal insurance problem as fuzzy random variable. Not only upper and lower bounds on premiums are derived but an uncertain optimal insurance problem is set up taking into account two sources of uncertainty. Abdullah and Rahman (2012) take advantage of a fuzzy inference system to study the likelihood of acquiring a health insurance. They investigate risk factors which might influence the possibility of buying a health insurance in order to identify potential customers (cf. Abdullah and Rahman 2012, p. 116).

Casanovas et al. (2015) provides a method to implement expert knowledge in the decision-making process referring to pricing. Not only information on cost structures and the actuarial premium is considered but the method also enables us to implement expert knowledge e.g. referring to the strategic vision.

The work Gençtürk et al. (2011) presents an application of fuzzy set theory for the pricing of high excess of loss layers in a reinsurance context. The authors aim to provide a fuzzy price which is an interval referring to "reasonable" prices instead of just a crisp premium. Therefore, the parameters of a general Pareto distribution are considered as fuzzy numbers (cf. Gençtürk et al. 2011, pp. 24f.).

3.3 Risk Classification

The technique of risk classification aims to distinguish insurable risks. Hereby, risks are categorized according to their severity (claim size) and their frequency (probability) of occurrence (cf. Ebanks et al. 1992). These classifications are needed for underwriting as well as for pricing purposes. If relevant findings with respect to an appreciate classification were ignored within the pricing this might lead to non-risk-adequate premiums and therefore potentially effects of anti-selection. Hence risk classification can be understood as an early fundamental step of actuarial analysis.

Ebanks et al. (1992) propose a fuzzy classification method. They consider an ideal risk as a point of reference. It is assumed that this risk can be described by a number of different risk characteristics. Now each other individual risk is assessed by regarding to which extent it corresponds to the risk characteristics of the ideal risk. This assessment is measured with the help of membership functions. Based on this, they present an approach to identify and evaluate sets of preferred risks with the help of three different measures of fuzziness (Additionally they gave an example from a life insurance context following Lemaire (1990)).

Ostaszewski (1993) illustrates that in the case of risk classification it is often not sensible to apply strict criteria. Some boundaries (of strict criteria) might be accumulated with "bad" risks. Therefore, there might be an incentive to terminate existing contracts and, consequently, there are phenomena of antiselection (cf. Ostaszewski 1993, p. 60). Thus, Lemaire (1990) and other publications move on to make vague statements used for risk classification. The use of fuzzy sets may be a good idea in this case as these aim to model imprecise human reasoning (cf. Ostaszewski 1993, p. 59).

Ostaszewski (1993) points out that classical approaches of classification are based on assumptions stating which are the relevant variables to describe the risk. This can lead to situations in which information – that is generally available – is not analyzed because it was not selected due to hazard or a missing corresponding intuition. Hence, he is interested in developing a less restrictive method that enables the actuary to detect the relevant pattern for classification (cf. Ostaszewski 1993, pp. 50ff.). His solution is based on the technique of the c-means algorithm (cf. Bezdek 1981). This algorithm is a fuzzy method of data clustering. It differs from crisp clustering approaches as it allows data elements to be part of several different clusters (each with a certain degree of membership, cf. Zimmermann et al. 1999, pp. 226ff.).

His concept (of the c-means algorithm) is applied by Derrig and Ostaszewski (1995) to a car insurance data set. Their objective is to define rating territories in order to group policies according to cities regional risk profiles. The classification itself is allowed to be fuzzy in the sense that a city could be "partial" member in different clusters.[31] Verrall and Yakoubov (1999) transfer the c-mean clustering method by Derrig and Ostaszewski (1995) to the problem of defining appropriate groupings by policyholder age. They give an extensive numerical example how this method could be used in practice.

Although Horgby (1998)'s work on risk classification refers to life insurance we would like to mention it as well because we are convinced that his method is transferable to the non-life sector. He shows in a three step approach how risks that are characterized by multiple fuzzy risk factors can be combined and, hence, classified. What makes his work interesting for use in practice is that his method is a fuzzy extension of a numerical rating approach. In a first step (fuzzification), various vague and imprecise information on the health status are modeled as different fuzzy sets (e.g. "mild level of cholesterol"). The next step is concerned with merging these fuzzy prognostic factors. Corresponding rules are exemplarily defined. This results in a fuzzy evaluation of necessary risk loadings. In his third step (defuzzification) he transfers this fuzzy number to a crisp premium loading. Consequently, his work could have been as well mentioned within the class of papers on pricing.

[31] A second part of their work deals with a fuzzy clustering based expert system to rank claims with respect to their potential being fraudulent.

Başer et al. (2011) apply an adaptive network based fuzzy inference system for risk classification in life insurance. Policyholders are classified by their cardiovascular risk characteristics, namely blood pressure, cholesterol level, obesity and smoking behavior. Also a comparison of the results of the proposed model using adaptive network models with results obtained from fuzzy regression is conducted.

3.4 Claims Reserving

A further conceivable and interesting area of application is without question the field of claims reserving. Claims reserving deals with the quantification of outstanding loss liabilities and, likewise, with the quantification of the liabilities' risk. An introduction to claims reserving as well as an overview of popular reserving methods is presented in Chapter 4.

Presenting a summary of applications of fuzzy methods in claims reserving is one of the main goals of this work. That is why we devote a separate section in the chapter dealing with claims reserving methods to the summary and the discussion of the fuzzy claims reserving methods (cf. Section 4.6). Before doing so, the most important claims reserving methods – both deterministic and stochastic – which are used in the progress of this work are presented in Chapter 4 (cf. Sections 4.1-4.5). The most important publications in this area are Andrés Sánchez and Terceño Gómez (2003), Andrés Sánchez (2006, 2007, 2012), and Başer and Apaydin (2010). The methods presented in Chapters 5, 6 and 7 can be classified in the same field of application and utilize similar approaches. The findings in Chapter 5 are based on the publication of Heberle and Thomas (2014) whereas the results in Chapter 7 are based on the publication of Heberle and Thomas (2016). The model presented in Chapter 6 is developed within the scope of this dissertation.

4 Methods of Claims Reserving

In the past years the field of claims reserving in the insurance industry has got a growing importance. Claims reserving refers to the prediction of outstanding loss liabilities for not finally settled insurance claims as well as the quantification of the corresponding risk and is an important field in non-life insurance mathematics. For this purpose, future insurance technical obligations need to be quantified and assessed. The claims reserve, i.e. the predicted outstanding loss liabilities, refers to a large position on the liabilities side of an insurance company's balance sheet. Hence, neither over- nor underestimating of the outstanding loss liabilities is appropriate. If they are overestimated, this would temporarily reduce the insurance company's earnings and thus its equity capital. Since the reserves can serve as a basis for pricing calculations, too high premiums might turn out. Likewise, the reserve needs to be high enough to settle all future loss liabilities. If it is not high enough, unexpected run-off losses might occur in later years and / or premiums for new business of an insurance company might be possibly too low.

The size of the outstanding loss liabilities is estimated with the help of mathematical models. On the one hand there are algorithmic models like e.g. the chain-ladder- (CL), Bornhuetter Ferguson- (BF) or the additive method. On the other hand outstanding loss liabilities can be e.g. estimated by distributional and Bayesian models as well as by generalized linear models (GLMs) and Bootstrap methods. Surveys on methods of claims reserving are e.g. given in Wüthrich and Merz (2008), Taylor (2000), Kaas et al. (2008, Chapter 10), and Mack (2002, Chapter 3). Those methods can be classified into distribution-free and distributional ones wherein the ones presented here belong to the former. A generalization of distribution-free reserving methods as linear stochastic reserving methods (LSRMs) is considered in Dahms (2012). Distributional methods are among others regarded in England and Verrall (2002) and Merz and Wüthrich (2010). In this context often not only the first two moments but a predictive distribution can be deduced. Moreover, the paid-incurred chain (PIC) claims reserving method also allows for the incorporation of two sources of information, i.e. paid and incurred data (cf. Merz and Wüthrich 2010).[32]

In this chapter notations and conventions are presented at first in Section 4.1. Subsequently, several common claims reserving methods are stated with a particular emphasis on those methods used in the proceeding of this work. It begins with a presentation of the CL-method in Section 4.3 and the BF-method in Section 4.4. Section 4.5 serves as a provision of those claims reserving methods which are used in the following Section 4.6. A survey on applications of fuzzy methods in claims reserving is given in Section 4.6 comprising a corresponding discussion in Section 4.6.7.

4.1 Notations and Definitions

An insurance policy is a contract between an insurer and the insured. The insured pays a premium to the insurer to assure against an uncertain loss. In case of a covered claim the insurer will compensate for it (at least partly). In this work we will only consider situations in a non-life insurance setting. The situation for life insurance contracts is different and is not subject of this dissertation. Many life insurance contracts are fixed-benefit insurances and their calculation differs from the non-life case.
At the end of a business year not all claims which occurred during this year are usually settled. For the outstanding loss liabilities of these not completely regulated claims a "suitable" reserve needs to be set up. We distinguish two types of claims for which the claims amount is not finally known (cf. Wüthrich and Merz 2008, p. 3):

- **IBNR-claims** (incurred but not reported): The claim has occurred but has not been reported yet. For these claims an IBNR-reserve needs to be set up as these claims might only be reported many years after they have taken place. As examples for IBNR-claims one can regard an engineer's fault in construction of a bridge which might be only realized during special wind conditions or when an earthquake

[32] As a further example for the incorporation of both paid and incurred data the work of Quarg and Mack (2004) can be mentioned.

takes place. Considering claims with bodily injuries there might be (psychological) long-term effects which might not be diagnosed at first sight.

- **IBNER-claims** (incurred but not enough reserved): The claim has been reported but the final claims amount is not certain yet. Therefore, the reserves might not be precisely specified and, particularly, possibly not large enough. An example for an IBNER-claim is a long lasting judicial proceeding as e.g. the Contergan law suit in the 1960's.

The field of claims reserving offers models and methods to predict outstanding liabilities. This includes the challenge of predicting the ultimate claims.

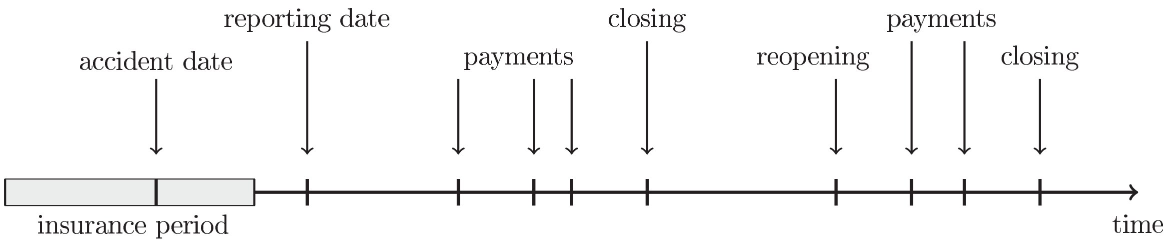


Figure 4.1: Typical development of a non-life insurance claim (adapted from Wüthrich and Merz 2008, p. 2).

A typical development of a claim is depicted in Figure 4.1. During the insurance period an accident (claim) happens to the insured. Depending on the type of loss it can take several years until the claim is reported (e.g. asbestos, fault in construction that is only obvious under certain circumstances, ...). Subject to the line of business (LoB) it can take some years until the claim is settled. Examples for LoBs are third party liability, Motor third party, health, fire, private and commercial property, ... Claims in property insurance are usually settled quickly whereas the settlement of liability claims might take a long period of time. A settlement does not necessarily need to be a final settlement. Referring to the example of the fault in construction at a bridge it might happen that the constructional flaw is not removed completely. Consequently, the claim needs to be reopened again and after another payment the claim is closed a second time (cf. Wüthrich and Merz 2008, p. 2).

Considering claims reserving, mostly common is the use of claims development trapezoids (or *run-off trapezoids*) (see Figure 4.2). By i we denote the relative accident year (year of occurrence of the claim) and j stands for the relative development year. Random

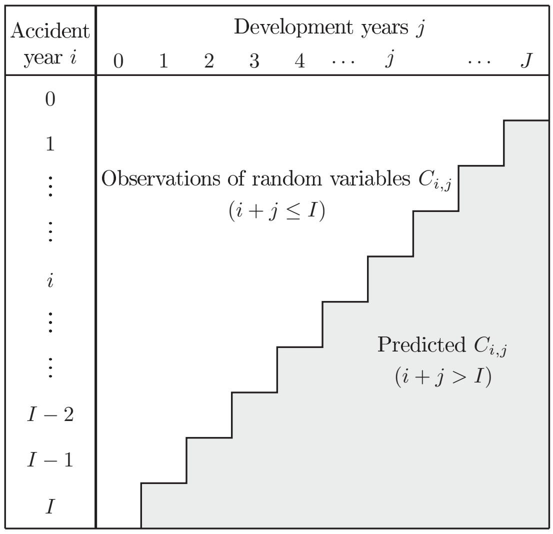


Figure 4.2: Claims development triangle (adapted from Wüthrich and Merz 2008, p. 10)

variables $C_{i,j}$ represent cumulative claims made in relative accident year i and relative development year j. We assume $i \in \{0, \dots, I\}$ and $j \in \{0, \dots J\}$ (cf. Wüthrich and Merz 2008, p. 11). The set of observations at time I, i.e. the upper left part of the triangle, is given by

$$\mathcal{D}_I := \{C_{i,j} \mid i + j \leq I, 0 \leq j \leq J\} \tag{4.1}$$

and the lower right part of the triangle needs to be predicted. We refer to $C_{i,J}$ as the ultimate claim for accident year i. Furthermore, it is assumed throughout this dissertation that the number of development years equals the number of accident years, i.e. $I = J$. Therefore, we are only considering claims development triangles throughout this work. It aims to simplify the notation even though all formulas can be extended to development trapezoids. Beyond that, we assume that claims are settled after J years. Otherwise so-called tail factors need to be taken into consideration. This is e.g. examined in Mack (1999) and Merz and Wüthrich (2013a).

The magnitudes in Figure 4.2 are not necessarily cumulative figures but it is also common to consider e.g. incremental data.[33] By

$$X_{i,j} := C_{i,j} - C_{i,j-1}$$

[33]Depending on what claims reserving method is being considered in the progress of this work the underlying data base can differ. Therefore, we will always explicitly state what figures are considered whether it is paid, incurred, cumulative, incremental, etc.

we denote increments which describe the variation from one development year to the next one in relative accident year i. They are related to the cumulative figures in the following way:

$$C_{i,j} = \sum_{k=0}^{j} X_{i,k}$$

Then at time I, the claims reserve for a single accident year for the outstanding loss liabilities, for accident year i is given by

$$R_i := C_{i,J} - C_{i,I-i}, \qquad i \in \{0, \dots, I\}. \tag{4.2}$$

Thus, the aggregated reserve is given by

$$R := \sum_{i=0}^{I} R_i.$$

At time I the outstanding loss liabilities are not observable and, thus, the reserve for a single accident year R_i as well as the aggregated reserve R are not observable and need to be predicted. We refer to their predictions as $\hat{R}_i$ or $\hat{R}$, respectively, and they are yielded by

$$\hat{R}_i := \hat{C}_{i,J} - C_{i,I-i}, \qquad i \in \{0, \dots, I\},$$

i.e. the difference of the predicted ultimate claim and the last observation, and

$$\hat{R} := \sum_{i=0}^{I} \hat{R}_i.$$

There exist numerous methods for the prediction of the reserves and a selection is presented in the following Sections 4.3-4.5. Depending on the data and the structure of the given information different methods are advisable.

4.2 Mean Square Error of Prediction

As elaborated on before, the task of claim reserving does not only lie in deriving predictions for the outstanding liabilities but also in the quantification of their risk. As the reserves often take on high amounts this is of vital interest.

One of the most popular risk measures is the (conditional) mean square error of prediction (MSEP). It can be stated as follows (cf. Wüthrich and Merz 2008, p. 33):

Definition 4.1 (Mean square error of prediction (MSEP))
Let X be a random variable, $\mathcal{D}_I$ a set of observations and $\hat{X}$ a $\mathcal{D}_I$-measurable predictor for X. Then, the conditional mean square error of prediction is defined as

$$\mathrm{MSEP}_{X|\mathcal{D}_I}(\hat{X}) := E\left[(\hat{X} - X)^2 \mid \mathcal{D}_I\right].$$

For the MSEP the following decomposition holds:

$$\mathrm{MSEP}_{\mathcal{D}_I}(\hat{X}) = \underbrace{\mathrm{Var}(X \mid \mathcal{D}_i)}_{\text{(conditional) process variance}} + \underbrace{\left(\hat{X} - E[X \mid \mathcal{D}_I]\right)^2}_{\text{(conditional) estimation error}} \tag{4.3}$$

In the decomposition in Equation (4.3) the process variance characterizes the variance within a stochastic model, i.e. pure randomness, and the estimation error addresses the uncertainty in the parameter estimation.

4.3 Chain-Ladder Method

The most widely used method in practice is the chain-ladder method. Due to its simplicity it is easy to implement and still delivers rather good results. According to Taylor (2000) it was first mentioned by Harnek (1966). Although the classical CL-method used to be a solely algorithmic reserving procedure it can be set into a stochastic framework. An overview of stochastic models for the CL-method is given in Hess and Schmidt (2002). The distribution-free derivation of the CL-method has been firstly presented by Mack (1993) and can be stated as follows (cf. Wüthrich and Merz 2008, p. 37).

Model Assumptions 4.2 (Distribution-free Chain-Ladder)
We assume for the cumulative claims $C_{i,j}$:

a) Cumulative claims $C_{i,j}$ of different accident years i are independent.

b) There exist development factors $f_0, \dots, f_{J-1} > 0$ and variance parameters $\sigma_0^2, \dots, \sigma_{J-1}^2 > 0$ such that

$$\begin{aligned} E\left[C_{i,j} \mid C_{i,0}, \dots, C_{i,j-1}\right] &= f_{j-1} C_{i,j-1} \\ \operatorname{Var}(C_{i,j} \mid C_{i,0}, \dots, C_{i,j-1}) &= \sigma_{j-1}^2 C_{i,j-1} \end{aligned} \tag{4.4}$$

holds true for all $i \in \{0, \dots, I\}$ and $j \in \{1, \dots, J\}$.

Remark 4.3

The development factors $f_0, \dots, f_{J-1}$ are also referred to as age-to-age- or chain-ladder-factors (CL-factors).

Under Model Assumptions 4.2 we can derive the following properties for the CL estimators.

Lemma 4.4

Assume that Model Assumptions 4.2 hold true. Then we have

$$E\left[C_{i,J} \mid \mathcal{D}_I\right] = E\left[C_{i,J} \mid C_{i,I-i}\right] = C_{i,I-i} f_{I-i} \cdots f_{J-1}$$

for all $i \in \{1, \dots, I\}$.

Proof. According to Equation (4.1) the set of observations is given by $\mathcal{D}_I = \{C_{i,j} \mid i+j \leq I, 0 \leq j \leq J\}$. With the independence assumption on different accident years and the tower property for conditional expectations we get:

$$\begin{aligned} E\left[C_{i,J} \mid \mathcal{D}_I\right] &= E\left[C_{i,J} \mid C_{i,0}, \dots, C_{i,I-i}\right] \\ &= E\left[\, E[C_{i,J} \mid C_{i,J-1}, \dots, C_{i,0}] \mid C_{i,0}, \dots, C_{i,I-i}\right] \\ &\overset{(4.4)}{=} E\left[f_{J-1} C_{i,J-1} \mid C_{i,0}, \dots, C_{i,I-i}\right] \\ &= f_{J-1}\, E\left[C_{i,J-1} \mid \mathcal{D}_I\right] \\ &= \dots \\ &= f_{J-1} \cdots f_{I-i}\, E\left[C_{i,I-i} \mid C_{i,0}, \dots, C_{i,I-i}\right] \\ &= C_{i,I-i} f_{I-i} \cdots f_{J-1} \end{aligned}$$

The proof is conducted analogously for $E\left[C_{i,J} \mid C_{i,I-i}\right]$. $\square$

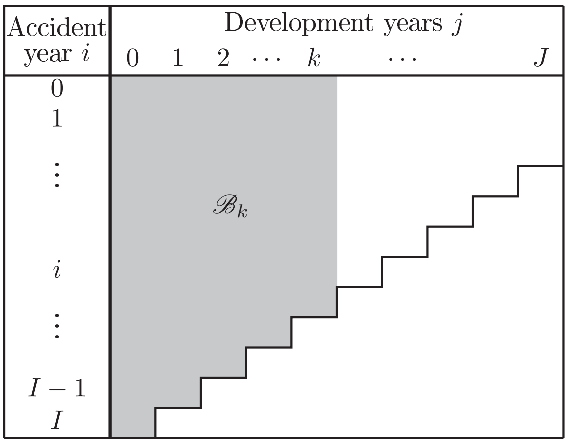


Figure 4.3: The set $\mathscr{B}_k$ of all observed cumulative claims up to development year k.

Therefore, the expected ultimate claim given the observations $\mathcal{D}_I$ is derived by successively multiplying the last observation for a given accident year with the CL-factors. Usually the CL-factors are not known. In these cases the unknown factors f_j need to be estimated as well. They are estimated by

$$\hat{f}_j := \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \frac{C_{i,j+1}}{C_{i,j}}, \qquad j \in \{0, \dots, J-1\}. \tag{4.5}$$

A predictor for the ultimate claim is hence given by

$$\hat{C}_{i,J} = C_{i,I-i} \hat{f}_{I-i} \cdots \hat{f}_{J-1}, \qquad i \in \{1, \dots, I\}. \tag{4.6}$$

The set of all observed cumulative claims up to development year k is given by $\mathscr{B}_k := \{C_{i,j} \mid i + j \leq I, 0 \leq j \leq k\} \subseteq \mathcal{D}_I$. An illustration of the set $\mathscr{B}_k$ is given in Figure 4.3. For $k = J$ we have $\mathscr{B}_J = \mathcal{D}_I$.

For the CL-estimators $\hat{f}_j$ the following properties can be derived.

Lemma 4.5
Let $\mathscr{B}_k := \{C_{i,j} \mid i + j \leq I, 0 \leq j \leq k\} \subseteq \mathcal{D}_I$. If Model Assumptions 4.2 hold true we have:

a) Given $\mathscr{B}_j$ the estimator $\hat{f}_j$ is unbiased, i.e. $E\left[\hat{f}_j \mid \mathscr{B}_j\right] = f_j$.

b) $\hat{f}_j$ is unconditionally unbiased, i.e. $E\left[\hat{f}_j\right] = f_j$.

c) *The estimators* $\hat{f}_0, \ldots, \hat{f}_{J-1}$ *are uncorrelated and it holds* $E\left[\hat{f}_0 \cdots \hat{f}_j\right] = E[\hat{f}_0] \cdots E[\hat{f}_j]$ *for all* $j \in \{0, \ldots, J-1\}$.

Proof. See Wüthrich and Merz (2008, p. 19). □

Remarks 4.6

a) The CL-estimators $\hat{f}_j$, $j \in \{0, \ldots, J-1\}$, are uncorrelated even though they partly take into account the same data.

b) The CL-estimator $\hat{f}_j$ is unbiased, i.e. on average it yields the true parameter.

Besides Mack's distribution-free chain-ladder model which puts the method into a stochastic framework there have been other attempts to establish a connection to statistical methods such as linear models. According to Brosius (1993) the idea of modeling age-to-age-factors with the help of regression techniques started off in the 1950's. Brosius (1993) referred to the development factors as link-ratios. Barnett and Zehnwirth (2000) argue that a link-ratio is a trend which can be modeled by a straight line through the origin (cf. Barnett and Zehnwirth 2000, p. 249). They show that several link-ratio techniques can be considered as weighted regression without intercept and introduce the "Extended Link Ratio Family" which comprises e.g. the CL and Cape Cod reserving methods.[34] A survey on the use of linear models in claims reserving is as well given in Ludwig et al. (2009).

The CL-factors can be understood as slope parameters of linear models. Precisely, the CL-estimators turn out as estimators for the slope in a weighted linear regression model without intercept, i.e. to determine the CL-factors a weighted linear regression model is considered for every factor.

For the estimation of the CL-factors $\hat{f}_j$, $j = 0 \ldots, J-1$, the following weighted linear model (cf. Definition A.2[35])

$$\mathbf{y}_j = \mathbf{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j$$

[34] Results for link-ratio techniques considered as a regression with or without intercept are also shown in Murphy (1994). A presentation of the Cape Cod method can be found e.g. in Wüthrich and Merz (2008, pp. 95 ff.).

[35] A short introduction and the notation concerning linear models used throughout this work are given in the appendix (see Appendix A).

where

$$\mathbf{y}_j := \begin{pmatrix} C_{0,j+1} \\ C_{1,j+1} \\ \vdots \\ C_{I-1-j,j+1} \end{pmatrix} \in \mathbb{R}^{(I-j)}, \quad \mathbf{X}_j := \begin{pmatrix} C_{0,j} \\ C_{1,j} \\ \vdots \\ C_{I-1-j,j} \end{pmatrix} \in \mathbb{R}^{(I-j)},$$

$$\mathbf{W}_j := \begin{pmatrix} C_{0,j} & 0 & \cdots & 0 \\ 0 & C_{1,j} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & C_{I-1-j,j} \end{pmatrix} \in \mathbb{R}^{(I-j)\times(I-j)} \quad \text{and} \quad \boldsymbol{\varepsilon}_j := \begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_{I-1-j} \end{pmatrix} \in \mathbb{R}^{(I-j)} \tag{4.7}$$

is regarded. In the further progress of this work we are considering cumulative claims such that the matrix $\mathbf{W}_j$ is positive definite. We yield the estimator for the regression coefficient $\boldsymbol{\beta}_j$ with the Aitken estimator (cf. Equation (A.1)):

$$\begin{aligned} \hat{\boldsymbol{\beta}}_j &= \left(\mathbf{X}_j^T \mathbf{W}_j^{-1} \mathbf{X}_j\right)^{-1} \mathbf{X}_j^T \mathbf{W}_j^{-1} \mathbf{y}_j \\ &= \left(\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix} \mathbf{X}_j\right)^{-1} \begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix} \mathbf{y}_j \\ &= \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} = \hat{f}_j, \quad j \in \{0, \ldots, J-1\} \end{aligned} \tag{4.8}$$

Comparing Equations (4.5) and (4.8) shows that both approaches lead to the same estimators for the CL-factors. By regarding

$$\mathbf{y}_j^* := \mathbf{W}_j^{-\frac{1}{2}} \mathbf{y}_j = \begin{pmatrix} \frac{C_{0,j+1}}{C_{0,j}^{1/2}} \\ \vdots \\ \frac{C_{I-1-j,j+1}}{C_{I-1-j,j}^{1/2}} \end{pmatrix}, \qquad \mathbf{X}_j^* := \mathbf{W}_j^{-\frac{1}{2}} \mathbf{X}_j = \begin{pmatrix} C_{0,j}^{1/2} \\ \vdots \\ C_{I-1-j}^{1/2} \end{pmatrix} \tag{4.9}$$

and

$$\boldsymbol{\varepsilon}_j^* := \mathbf{W}_j^{-\frac{1}{2}} \boldsymbol{\varepsilon}_j = \begin{pmatrix} \frac{\varepsilon_0}{C_{0,j}^{1/2}} \\ \vdots \\ \frac{\varepsilon_{I-1-j}}{C_{I-1-j,j}^{1/2}} \end{pmatrix} \tag{4.10}$$

the resulting model

$$\mathbf{y}_j^* = \mathbf{X}_j^* \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j^* \tag{4.11}$$

fulfills the requirements of a multiple linear model (cf. Appendix A). That is the reason why the CL-factors can also be derived with OLS when considering the transformed linear model (4.11). Thus, the Gauß-Markov theorem yields that the resulting estimator possesses the BLU-property. In order to determine all CL-factors for a given development triangle a sequence of general multiple linear models needs to be fitted to the data.

The CL-method has been considered in various scientific publications: the corresponding mean square error of prediction is e.g. subject in Mack (1993) as well as in Buchwalder et al. (2006). As an extension of the model, the additional incorporation of a priori information with the help of a credibility approach is regarded in Gisler and Wüthrich (2008). A multivariate version, i.e. a version in which several portfolios are examined simultaneously, is considered e.g. in Pröhl and Schmidt (2005) and the corresponding prediction error for the multivariate model is addressed in Merz and Wüthrich (2008).

Remark 4.7
The CL method has a number of advantages since it is distribution-free, easy to implement and can be interpreted intuitively. Moreover, it can be stated as a linear model and, thus, methods of well examined linear regression can be used.
As disadvantage it has to be stated that only the last observation for a given accident year is taken into account to predict the ultimate claims and, thus, the reserve (cf. Equation (4.6)). Consequently, the CL method is sensitive to outliers or zeros on the diagonal. In addition, the homogeneity characteristics should be fulfilled, i.e. among others no trend in the development factors should be existent.

In contrast to the CL-method, the BF-method presented in the next section can also incorporate additional information. Subsequent to a definition both methods are compared.

4.4 Bornhuetter Ferguson Method

The BF method has also been introduced as a purely computational reserving method which can comprise a priori information in the form of external and/or internal information as well. The a priori information derives from market statistics, business plans, expert knowledge, organizational data, etc. The method has been firstly presented by Bornhuetter and Ferguson (cf. Bornhuetter and Ferguson 1972). Summaries are given e.g. in Wüthrich and Merz (2008, pp. 21ff.), Schmidt and Radtke (2004, pp. 37ff.) and Mack (2000). The development triangle and the given a priori information are presented in Figure 4.4.

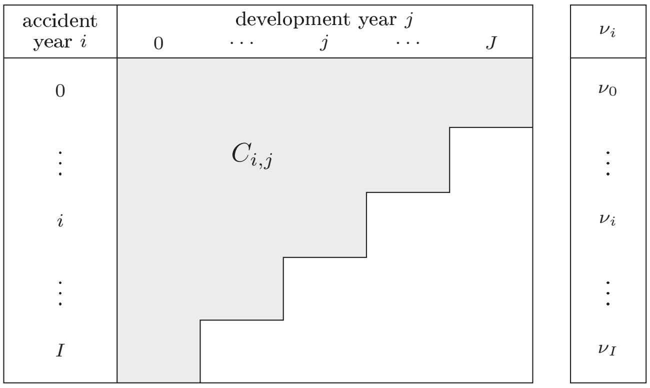


Figure 4.4: Development triangle at time $t = I$ with observable cumulative claims $C_{i,j}$ in the upper left part and given a priori information ν_i deriving from e.g. expert knowledge on the right.

As before, the upper left part of the triangle contains the observations at time I and the lower right part needs to be predicted. The right column in Figure 4.4 represents the a priori information ν_i which are a priori estimates for the ultimate claims.

The BF method can also be put into a stochastic framework and there are multiple ways to formalize it (cf. e.g. Mack (2008) or Verrall (2004)). We stick to the following definition (cf. Wüthrich and Merz 2008, p. 21).

Model Assumptions 4.8 (Bornhuetter Ferguson (BF) method)
We assume for cumulative claims $C_{i,j}$, $i \in \{0, \dots, I\}$, $j \in \{0, \dots, J\}$:

a) Cumulative claims $C_{i,j}$ of different accident years i are independent.

b) There exist parameters $\nu_0, \dots, \nu_I > 0$ and a pattern $\gamma_0, \dots, \gamma_J > 0$ with $\gamma_J = 1$ such that for all $i \in \{0, \dots, I\}$, $j \in \{0, \dots, J-1\}$ and $k \in \{1, \dots, J-j\}$ we have:

$$E[C_{i,0}] = \gamma_0 \nu_i$$
$$E[C_{i,j+k} \mid C_{i,0}, \dots, C_{i,j}] = C_{i,j} + (\gamma_{j+k} - \gamma_j)\, \nu_i$$

With Model Assumptions 4.8 we yield

$$E[C_{i,j}] = \gamma_j \nu_i \qquad \text{and} \qquad E[C_{i,J}] = \nu_i. \tag{4.12}$$

In particular, the expected ultimate claim $E[C_{i,J}]$ for accident year i is given by the a priori information ν_i. The sequence $(\gamma_j)_{j\in\{0,\dots,J\}}$ is referred to as claims development pattern and for cumulative claims payments it is called payout pattern. Moreover, we speak of the factors $(1-\gamma_{I-i})$ as still-to-come factors which specify the share of claims not being settled yet. In addition, we can conclude

$$E[C_{i,I} \mid C_{i,0}, \dots, C_{i,I-i}] = C_{i,I-i} + (1-\gamma_{I-i})\nu_i.$$

This observation motivates the following predictor for ultimate claims $C_{i,J}$ given by

$$\widehat{C_{i,J}}^{\text{BF}} = C_{i,I-i} + (1-\hat{\gamma}_{I-i})\,\hat{\nu}_i \qquad \text{for all } i \in 1,\dots,I. \tag{4.13}$$

From Equation (4.13) we see that a higher weight is put on the a priori estimates ν_i for earlier development years whereas the influence for later development years is decreasing. Moreover, if both a priori estimates $\hat{\gamma}_{I-i}$ and $\hat{\nu}_i$ originate externally, the predictor only depends on the last observation and does not use any more information from the development triangle.

We can derive a predictor for the outstanding claims from Equation (4.13) as

$$\widehat{R}_i^{\text{BF}} = (1-\hat{\gamma}_{I-i})\,\hat{\nu}_i \qquad \text{for all } i \in 1,\dots,I. \tag{4.14}$$

In Equations (4.13) and (4.14) the a priori estimates $\hat{\gamma}_{I-i}$ and $\hat{\nu}_i$ can be both given externally. Then, the BF method is applied in its pure form as it has been intended. These estimates originate e.g. from plan values of business plans or expert knowledge and need to be given in advance. In practice, actuaries sometimes diverge from the BF method in its pure form and derive estimates with the help of CL-estimators, i.e. information given in the development triangle. Then, we observe the following.

Model Assumptions 4.2 yield with the tower property for conditional expectations

$$E[C_{i,j}] = E[\,E[C_{i,j} \mid C_{i,j-1}]] = E[f_{j-1}C_{i,j-1}] = f_{j-1}\,E[C_{i,j-1}] = \dots = E[C_{i,0}]\prod_{k=0}^{j-1} f_k$$

and

$$\begin{aligned} E[C_{i,J}] &= E[C_{i,0}] \prod_{k=0}^{J-1} f_k \\ \iff E[C_{i,0}] &= \prod_{k=0}^{J-1} f_k^{-1}\, E[C_{i,J}]. \end{aligned}$$

Thus, we have the following representation of an expected cumulative claim in the context of the CL method:

$$E[C_{i,j}] = \prod_{k=j}^{J-1} f_k^{-1}\, E[C_{i,J}] \tag{4.15}$$

There is a similarity between the CL and BF method since in Equation (4.15) the factor $\prod_{k=j}^{J-1} f_k^{-1}$ has taken the place of γ_j (cf. Equation (4.12)). Consequently, we can deduce a sequence $(\gamma_j)_{j \in \{j,\ldots,J\}}$ if the CL-factors are known and vice versa. Thus, the factors γ_j, $j \in \{0,\ldots,J\}$ can be estimated from the development triangle by

$$\hat{\gamma}_j^{\mathrm{CL}} := \hat{\gamma}_j := \prod_{k=j}^{J-1} \hat{f}_k^{-1}. \tag{4.16}$$

Applying the estimator $\hat{\gamma}_j^{\mathrm{CL}}$ for γ_j leads to the following predictor for the ultimate claim in the BF method:

$$\widehat{C_{i,J}}^{\mathrm{BF}} = C_{i,I-i} + \left(1 - \hat{\gamma}_{I-i}^{\mathrm{CL}}\right) \hat{\nu}_i$$

The predictor for the ultimate claim in the CL method can be represented as (cf. Equation (4.6))

$$\begin{aligned} \widehat{C_{i,J}}^{\mathrm{CL}} &= C_{i,I-i} \prod_{j=I-i}^{J-1} \hat{f}_j \\ &= C_{i,I-i} + C_{i,I-i} \left(\prod_{j=I-i}^{J-1} \hat{f}_j - 1 \right) \\ &= C_{i,I-i} + \frac{\widehat{C_{i,J}}^{\mathrm{CL}}}{\prod_{j=I-i}^{J-1} \hat{f}_j} \left(\prod_{j=I-i}^{J-1} \hat{f}_j - 1 \right) \\ &= C_{i,I-i} + \left(1 - \frac{1}{\prod_{j=I-i}^{J-1} \hat{f}_j} \right) \widehat{C_{i,J}}^{\mathrm{CL}}. \end{aligned}$$

Thus, as a difference between the predictors for the ultimate claims in the BF and CL method we see that the BF method uses an a priori information $\hat{\nu}_i$ whereas the CL method applies the prediction $\widehat{C_{i,J}}^{\mathrm{CL}}$. Consequently, we can identify the claims development patterns of the BF and CL method when taking the estimates $\hat{\gamma}_j^{\mathrm{CL}}$ as given in Equation (4.16). The expected ultimate claims $E[C_{i,J}]$ for both methods differ since the BF method relies on an a priori information $\hat{\nu}_i$ and the prediction $\widehat{C_{i,J}}^{\mathrm{CL}}$ in the CL method is based solely on observations.

The mean square error of prediction of the BF method in a generalized linear model (GLM) framework using an overdispersed Poisson model is presented in Alai et al. (2009) and Alai et al. (2011). It has been investigated in a more general context concerning the underlying idea of the BF method in Saluz et al. (2011).

Remark 4.9
One advantage of the BF method is that it not only provides a mean to comprise a priori information in a claims reserving context but is also very robust. Depending on the available data there might be instabilities in the settlement of earlier development years. The use of a priori estimates ν_i of the expected ultimate claims counteracts these instabilities.
However, the BF method does not use all available information since it only applies the claims made in calendar year I, i.e. the observations $C_{i,I-i}$ on the last observed diagonal in the development triangle $\mathcal{D}_I$.

4.5 Other Claims Reserving Methods

In this section a number of other distribution-free claims reserving methods are presented. We focus on those methods that have been used in combination with fuzzy theory so far (see Section 3.4). In this way, this section is not intended to be an extensive overview of distribution-free claims reserving methods.

4.5.1 Claims Reserving with London Chain Approach

In their 1986 work Benjamin and Eagles investigated claims reserving with a special focus on Lloyd's and the London market (cf. Benjamin and Eagles 1986)). They pursuit a regression approach by adapting straight lines to the given data. This idea is also seized and discussed in Straub (1988, pp. 109ff.). Andrés Sánchez and Terceño Gómez (2003) tie into this suggestion with a fuzzy approach (see Section 4.6.1).

Benjamin and Eagles (1986) regard paid loss ratios, i.e. claims payments set in proportion to earned premiums. In order to predict the ultimate loss ratio they plot all observed paid loss ratios for several development years of a given accident year. The authors fit a line to the loss ratios and, thereby, presume a linear relationship (cf. Benjamin and Eagles 1986, pp. 214ff.).

The idea of considering a linear relationship is picked up by Straub and referred to as the "London Chain" approach (cf. Straub 1988, p. 110). Let $Z_{i,j}$, $i, j \in \{0, \dots, I\}, i+j \leq I$, be claims payments for claims occurred in accident year i and development year j. Then, the following linear relationship for paid claims for each development year is assumed:

$$Z_{i,j+1} = a_j + b_j Z_{i,j}, \qquad i \in \{0, \dots, I-j-1\}, j \in \{0, \dots, I-1\}$$

where $a_j, b_j \in \mathbb{R}$. The coefficients a_j and b_j are then determined with a least squares approach, i.e.

$$\arg\min_{a_j, b_j} \sum_{i=0}^{I-j-1} \left(Z_{i,j+1} - a_j - b_j Z_{i,j}\right)^2 .$$

Hence, the coefficients are specified in the way that the quadratic distance is minimized. We refer to the estimates of the coefficients as $\hat{a}_j$ and $\hat{b}_j$. Thus, the ultimate claim for each accident year i can be predicted by

$$\hat{Z}_{i,J} = \hat{a}_{J-1} + \hat{b}_{J-1}\Big(\dots \hat{a}_{j+2} + \hat{b}_{j+2}\big(\hat{a}_{j+1} + \hat{b}_{j+1}(\hat{a}_j + \hat{b}_j Z_{i,j})\big)\Big). \tag{4.17}$$

4.5.2 Claims Reserving with a Smoothing Approach by Sherman

The method to which Andrés Sánchez applied fuzzy regression methods makes use of individual development factors and a smoothing approach presented in Sherman (1984). According to Andrés Sánchez (2006) the method is conducted in four steps. As a first step, the individual claims development factors are calculated, i.e.

accident-year	development year j						
	0	1	$\cdots$	j	$\cdots$	$J-1$	J
0	$F_{0,0} := \frac{C_{0,1}}{C_{0,0}}$	$F_{0,1} := \frac{C_{0,2}}{C_{0,1}}$	$\cdots$	$F_{0,j} := \frac{C_{0,j+1}}{C_{0,j}}$	$\cdots$	$F_{0,J-1} := \frac{C_{0,J}}{C_{0,J-1}}$	
1	$F_{1,0} := \frac{C_{1,1}}{C_{1,0}}$	$F_{1,1} := \frac{C_{1,2}}{C_{1,1}}$	$\cdots$	$F_{1,j} := \frac{C_{1,j+1}}{C_{1,j}}$	$\cdots$		
$\vdots$	$\vdots$				$\iddots$		
i	$F_{i,0} := \frac{C_{i,1}}{C_{i,0}}$	$F_{i,1} := \frac{C_{i,2}}{C_{i,1}}$	$\cdots$				
$\vdots$	$\vdots$		$\iddots$				
$I-1$	$F_{I-1,0} := \frac{C_{I-1,1}}{C_{I-1,0}}$						
I							

Table 4.1: Claims development triangle with individual development factors $F_{i,j} := \frac{C_{i,j+1}}{C_{i,j}}$

$$F_{i,j} := \frac{C_{i,j+1}}{C_{i,j}}, \qquad j \in \{0, \ldots, J-1\}, i \in \{0, \ldots, I-j-1\},$$

as shown in Table 4.1.

In a second step, a representative development factor for each development year j, $j \in \{0, \ldots, J-1\}$, is chosen. The author refers to the Claims Reserving Manual by the Institute of Actuaries in which taking an average of the factors is proposed (cf. Institute of Actuaries 1997, p. B.7.1). Therefore, we calculate for each development year j

$$\overline{f}_j := \frac{1}{I-j} \sum_{i=0}^{I-j-1} F_{i,j}, \qquad j \in \{0, \ldots, J-1\}.$$

As a next step, the author relies on Sherman and conducts a similar approach (cf. Sherman 1984, p. 126). In order to reduce the complexity of the estimation problem a curve is fitted to the development factors $\overline{f}_j$:

$$\begin{aligned} & \overline{f}_j = 1 + e^a (j+1)^b \\ \Leftrightarrow \qquad & \overline{f}_j - 1 = e^a (j+1)^b, \qquad j \in \{0, \ldots, J-1\}, a, b \in \mathbb{R} \end{aligned} \tag{4.18}$$

By doing so, only two instead of J parameters need to be estimated. Equation (4.18) can be conveyed into a linear equation by taking logarithms:

$$\begin{aligned} & \ln\left(\overline{f}_j - 1\right) = a + b\ln(j+1) \\ \Leftrightarrow \qquad & F^{\log} = a + b\ln(j+1), \qquad j \in \{0,\ldots,J-1\}, \end{aligned} \tag{4.19}$$

where $F^{\log} := \ln\left(\overline{f}_j - 1\right)$. In order to employ OLS estimation techniques (see Appendix A) an error term ε_j is added, $\varepsilon_j \sim \mathcal{N}(0,\sigma^2)$, $\sigma^2 \in \mathbb{R}_+$, i.e. we regard

$$F_j^{\log} = a + b\ln(j+1) + \varepsilon_j, \qquad j \in \{0,\ldots,J-1\}.$$

With the help of OLS estimation methods the parameters a and b are estimated where $F_j^{\log}$ is taken as dependent variable and $\ln(j+1)$ is the independent variable, $j \in \{0,\ldots,J-1\}$. Subsequently, a retransformation to the magnitudes considered at the beginning using Equations (4.19) and (4.18) leads to estimates for the development factors. Thus, the unknown values in the lower right part of the development triangle can be predicted and the reserves can be calculated.

4.5.3 Taylor's Geometric Separation Method

The geometric separation method dates back to Taylor (1977, 1978) and assumes a multiplicative structure for the entries of the claims development triangle.[36]

Let $X_{i,j}$, $i,j \in \{0,\ldots,I\}$, $i+j \le I$, be incremental claims payments for claims occurred in accident year i and paid in development year j in this section. By N_i we denote the number of claims occurred in accident year i. Then, $S_{i,j} := \frac{X_{i,j}}{N_i}$ describes average incremental payments per claim for accident year i and development year j. The geometric separation method assumes that $S_{i,j}$ can be separated into two factors, i.e.

$$S_{i,j} \approx P_j \lambda_{i+j}, \tag{4.20}$$

where P_j, $\lambda_{i+j} \in \mathbb{R}$. Thus, the average incremental payments per claim in a claims development triangle can be split into a factor P_j standing for the share of claims payments settled in development year j and a second factor λ_{i+j} describing calendar

[36]Separation models can be classified in geometric and arithmetic separation models. The latter presumes an additive structure and has been studied e.g. in Taylor (1978). We will restrict our analysis to the geometric one.

year effects, i.e. effects on the diagonal of the triangle (cf. Kaas et al. 2008, p. 273). When additionally assuming that $\sum_{j=0}^{J} P_j = 1$ the parameters can be estimated with the help of linear regression techniques. We consider the logarithm of Equation (4.20) and assume an additional error term to yield a linear structure:

$$\ln S_{i,j} = \ln P_j + \ln \lambda_{i+j} + \varepsilon_{i,j}, \quad i, j \in \{0, \ldots, J\}, i + j \leq J, \tag{4.21}$$

where $\varepsilon_{i,j}$ is an error term. The linear model in Equation (4.21) can be stated in matrix notation as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{4.22}$$

where

$$\begin{aligned}
\mathbf{Y} &:= (\ln S_{0,0}, \ldots, \ln S_{0,J}, \ln S_{1,0}, \ldots, \ln S_{1,J-1}, \ldots, \ln S_{I,0})^T \\
\boldsymbol{\beta} &:= (\ln P_0, \ldots, \ln P_J, \ln \lambda_0, \ldots, \ln \lambda_J)^T \\
\boldsymbol{\varepsilon} &:= (\ln \varepsilon_{0,0}, \ldots, \ln \varepsilon_{0,J}, \ln \varepsilon_{1,0}, \ldots, \ln \varepsilon_{1,J-1}, \ldots, \ln \varepsilon_{I,0})^T
\end{aligned}$$

and

$$\mathbf{X} := \begin{pmatrix}
\multicolumn{2}{c}{\mathbf{I}_{(J+1)}} & \multicolumn{2}{c}{\mathbf{I}_{(J+1)}} \\
\mathbf{I}_J & \mathbf{0}_{J\times 1} & \mathbf{0}_{J\times 1} & \mathbf{I}_J \\
\vdots & \vdots & \vdots & \vdots \\
\mathbf{I}_{J-i+1} & \mathbf{0}_{(J-i+1)\times i} & \mathbf{0}_{(J-i+1)\times i} & \mathbf{I}_{J-i+1} \\
\vdots & \vdots & \vdots & \vdots \\
\mathbf{I}_1 & \mathbf{0}_{1\times J} & \mathbf{0}_{1\times J} & \mathbf{I}_1
\end{pmatrix}.$$

Here $\mathbf{I}_n$, $n \in \{1, \ldots, J\}$, denotes the identity matrix with n rows and columns and $\mathbf{0}_{k\times l}$, $k, l \in \{1, \ldots, J\}$, indicates a zero matrix with k rows and l columns. Consequently, the vector $\mathbf{Y}$ contains $\frac{1}{2}(J+1)(J+2)$ elements and $2J+2$ parameters in the parameter vector $\boldsymbol{\beta}$ need to be estimated. The parameter vector can then be estimated with OLS (see Appendix A). However, the linear model given in Equation (4.22) does not lead to a unique solution as $\mathbf{X}^T\mathbf{X}$ is singular. In order to overcome this problem, one fixes one parameter by setting it to a fix value and, consequently, leaves out one column in the design matrix $\mathbf{X}$ (cf. Goovaerts et al. 1990, p. 274; Taylor 1978, p. 49). Thus, the remaining $2J+1$ parameters need to be estimated.

4.5.4 Reserving with Kremer's Two Way Model of ANOVA

Kremer (1982) proposes a model which aims to take advantage of results of linear models, especially the analysis of variance (ANOVA). The model emanates from a multiplicative structure similar to those of the geometric separation method (see Section 4.5.3).

As before, $C_{i,j}$ denotes cumulative claims payments made in accident year i and development year j. Here, incremental claims payments $X_{i,j}$ are considered. We assume that the incremental claims payments are strictly positive, i.e. $X_{i,j} > 0$ for all $i, j \in \{0, \ldots, I\}, i + j \leq I$.[37] The underlying model for incremental claims payments is stated as follows:

$$X_{i,j} = E_i p_j e_{i,j} \tag{4.23}$$

where $E_i := E[C_{i,J}]$ is the expected ultimate claim for accident year i, p_j denotes the expected share describing the increase of expected claims amounts in development year j and $e_{i,j}$ refers to a random variable with mean one, i.e. $E[e_{i,j}] = 1$. Additionally, it is required that all factors p_j add up to one over all development years, i.e. $\sum_{j=0}^{J} p_j = 1$.

In order to yield a linear structure to employ results of linear models the logarithm of Equation 4.23 is considered. Therefore, we obtain

$$\ln X_{i,j} = \eta + a_i + b_j + \varepsilon_{i,j} \tag{4.24}$$

where $\varepsilon_{i,j} := \ln e_{i,j}$, $a_i := \ln E_i - \frac{1}{n}\sum_{k=0}^{J} \ln E_k$, $b_j := \ln p_j - \frac{1}{n}\sum_{k=0}^{J} \ln p_k$ and $\eta := \frac{1}{n}\sum_{k=0}^{J} (\ln E_k + \ln p_k)$. In addition, it is assumed that the random variables $\varepsilon_{i,j}$, $i + j \leq I$, are uncorrelated with mean zero and a constant variance, i.e. $E[\varepsilon_{i,j}] = 0$ and $\mathrm{Var}(\varepsilon_{i,j}) = \sigma^2 \in (0, \infty)$. With these assumptions estimators for the parameters η, a_i and b_j can be derived in a linear regression context with the help of the Gauß-Markov theorem (cf. Kremer 1982, pp. 49 f.).[38] In order to get predictions for the ultimate claims and, thus, the reserves, the transformations made above need to be reversed.

[37]The assumption of strictly positive incremental claims is not a limitation to the model. In case of negative increments, a constant $x \in \mathbb{R}$ is chosen with $x < \min\limits_{\substack{(i,j)\in\{0,\ldots,I\}\times\{0,\ldots,J\} \\ i+j\leq I}} X_{i,j}$. Then, we consider strictly positive incremental claims shifted by the constant x, i.e. $X'_{i,j} := X_{i,j} - x$ (cf. Kremer 1982, p. 48).

[38]The assumptions made here are chosen in the way that they resemble those of the two-way model of ANOVA. We will concentrate on the model stated in Equation (4.24) as it is sufficient for the further analysis in our case.

4.6 Applications of Fuzzy Methods

While first applications of fuzzy set theory in actuarial and insurance science date back to the early 1980's (see Section 3.1) there are only a few publications in the field of claims reserving. To our best knowledge, the first article has been published in the year 2003 (cf. Andrés Sánchez and Terceño Gómez 2003). This section aims to give an overview of the published fuzzy claims reserving methods up to present.

Models in claims reserving can be commonly divided into

- purely computational (cf. e.g. CL and BF method in their original form)
- stochastic (cf. e.g. CL, BF, PIC, LSRM, ...) and
- Bayesian/Credibility (cf. e.g. Bayes CL, Conditional LSRM), which are a particular subset of stochastic methods,

models.[39] Stochastic and Bayesian models have in common that they model stochastic randomness whereas Bayesian and Credibility models also allow to incorporate external data e.g. if the data is scarce.

All methods presented in this section are applications of fuzzy set theory and, thus, model vagueness (cf. Section 2.1.4). According to the current state of research models can be classified in the following way: While to our knowledge earlier models are all based on the application of FR (cf. e.g. Andrés Sánchez and Terceño Gómez 2003; Andrés Sánchez 2006, 2007; Başer and Apaydin 2010) the fuzzy chain-ladder and the fuzzy Bornhuetter Ferguson method comprise FNs (cf. Chapters 5 and 7). The reserving method presented in Chapter 6 is an application of FR as well and is on the basis of the CL method. In the following we will present the applications of FR in claims reserving in chronological order. Subsequently, we will discuss the methods regarding the appropriateness in claims reserving.

[39] This classification is not exhaustive since claims reserving methods can e.g. also be considered in a time series context or with applications of the Kálmán filter. Nonetheless, we stick to the rougher classification as it is sufficient in our case.

4.6.1 Fuzzy London Chain-Ladder Method by Andrés Sánchez and Terceño Gómez

Among the first utilizations of fuzzy methods has been the approach by Andrés Sánchez and Terceño Gómez to apply a FR method by Tanaka and Ishibuchi (1992) to a claims reserving method described in Straub (1988, pp. 109f.) which has been motivated by Benjamin and Eagles (1986) (see Section 4.5.1).[40] The regression model presented in Tanaka and Ishibuchi (1992) is the same as the Tanaka (1987) model as presented in Equations (2.13a)-(2.13d) (see p. 52).

The authors mention as their incitement to incorporate FR that OLS regression is only suitable if there is a good data basis available in the sense that there are enough observations to conduct OLS. Depending on the LoB this is not necessarily the case in claims reserving. Therefore, the authors propose the use of FR.

Andrés Sánchez and Terceño Gómez assume that a cumulative claim $\tilde{C}_{i,j+1}$ is a STFN, i.e. $\tilde{C}_{i,j+1} := (C_{i,j+1}, l_{C_{i,j+1}})$, and can be represented as

$$\tilde{C}_{i,j+1} = \tilde{a}_j \oplus \tilde{b}_j C_{i,j}, \qquad j = 0, \ldots, J-1$$

where $\tilde{a}_j$ and $\tilde{b}_j$ are STFN of the form $\tilde{a}_j := (a_j, l_{a_j})$ and $\tilde{b}_j := (b_j, l_{b_j})$. Then, a cumulative claim can be written as

$$\begin{aligned} \tilde{C}_{i,j+1} &= (C_{i,j+1}, l_{C_{i,j+1}}) \\ &= (a_j, l_{a_j}) \oplus (b_j, l_{b_j}) C_{i,j} \\ &= (a_j + b_j C_{i,j}, l_{a_j} + l_{b_j} C_{i,j}). \end{aligned} \tag{4.25}$$

As the FNs $\tilde{a}_j$ and $\tilde{b}_j$ are not observable in general they need to be estimated. We will denote the corresponding estimators by $\hat{\tilde{a}}_j = (\hat{a}_j, \hat{l}_{a_j})$ and $\hat{\tilde{b}}_j = (\hat{b}_j, \hat{l}_{b_j})$, respectively. In analogy to the non-fuzzy reserving method the estimates of the ultimate claims $\tilde{C}_{i,J}$, $i \in \{0, \ldots, I\}$, can be denoted by (cf. Equation (4.17))

$$\hat{\tilde{C}}_{i,J} = \hat{\tilde{a}}_{J-1} \oplus \hat{\tilde{b}}_{J-1}\Big(\ldots \hat{\tilde{a}}_{I-i+2} \oplus \hat{\tilde{b}}_{I-i+2}\Big(\hat{\tilde{a}}_{I-i+1} \oplus \hat{\tilde{b}}_{I-i+1}\big(\hat{\tilde{a}}_{I-i} \oplus \hat{\tilde{b}}_{I-i} C_{i,I-i}\big)\Big)\Big).$$

[40] The main focus of the publication of Andrés Sánchez and Terceño Gómez (2003) lies on the estimation of the term structure of interest. However, they mention further applications of FR in actuarial science, among which is the considered claims reserving method.

The reserve for each accident year i, $i \in \{0, \ldots, I\}$, is then given by

$$\tilde{R}_i = \tilde{C}_{i,J} \ominus C_{i,I-i}$$

and we yield a predictor for the reserve by

$$\hat{\tilde{R}}_i = \hat{\tilde{C}}_{i,J} \ominus C_{i,I-i}.$$

Analogously to the crisp case, the total reserve $\tilde{R}$ is given by $\tilde{R} = \bigoplus_{i=0}^{I} \tilde{R}_i$. The parameters are estimated by solving the resulting optimization problem yielded by FR. Here, the total spreads of the fuzzy output is minimized under certain constraints (cf. (2.13a)-(2.13d)). Considering (4.25) the parameters $\hat{\tilde{a}}_j$ and $\hat{\tilde{b}}_j$, $j \in \{0, \ldots, J-1\}$, can be determined with the following linear problem. Setting up the linear problem in this way acts on the assumption that all cumulative claims $C_{i,j}$ are positive:

$$\min \sum_{i=0}^{I-1} \left(l_{a_j} + l_{b_j} C_{i,j} \right) \tag{4.26a}$$

s.t.

$$a_j + b_j C_{i,j} - (l_{a_j} + l_{b_j} C_{i,j})(1 - \alpha^*) \leq C_{i,j+1} \qquad i = 1, \ldots, I-j \tag{4.26b}$$

$$a_j + b_j C_{i,j} + (l_{a_j} + l_{b_j} C_{i,j})(1 - \alpha^*) \geq C_{i,j+1} \qquad i = 1, \ldots, I-j \tag{4.26c}$$

$$l_{a_j}, l_{b_j} \geq 0 \tag{4.26d}$$

Equations (4.26b) and (4.26c) assure that the observed values fall into the estimated values. The above linear program needs to be solved J times to derive all development parameters $\hat{\tilde{a}}_j$ and $\hat{\tilde{b}}_j$. Equation (4.26d) ensures that the spreads are non-negative.

4.6.2 Fuzzy Version of Sherman's Scheme by Andrés Sánchez

A further application of FR methods is conducted in Andrés Sánchez (2006) to a reserving scheme proposed by Sherman (1984). In this approach development factors are being determined with the help of fuzzy regression techniques as well. The author claims as motivation that OLS only yields reasonable estimations if the data base is sufficiently large. As this usually might not be the case (depending on the LoB) he suggests the use of FR. In particular, he chooses the method mentioned in Ishibuchi and Nii (2001) (cf. pp. 53ff.). The approach is geared to the procedure as described in Section 4.5.2.

Firstly, for cumulative claims in the development triangle individual development factors are computed (cf. Table 4.1). For each development year $j \in \{0, \dots, J-1\}$ we consider the set of individual development factors $\{F_{0,j}, \dots, F_{I-j-1,j}\}$. Instead of taking the average of the individual development factors for a development year the maximum and minimum is considered. In analogy to the non-fuzzy case we define $F_{i,j}^{\log} := \ln(F_{i,j} - 1)$ and, thus,

$$\begin{aligned}\overline{F}_j &:= \ln\left(\max(F_{0,j}, F_{1,j}, \dots, F_{I-j-1,j}) - 1\right)\\ \underline{F}_j &:= \ln\left(\min(F_{0,j}, F_{1,j}, \dots, F_{I-j-1,j}) - 1\right), \qquad j \in \{0, \dots, J-1\}.\end{aligned}$$

Hence, the following FR model is considered

$$\tilde{F}_j := (F_j, l_{F_j}, r_{F_j}) = \tilde{a} \oplus \tilde{b} \ln(j+1)$$

where $\tilde{a}$ and $\tilde{b}$ are TFNs of the form $\tilde{a} := (a, l_a, r_a)$ and $\tilde{b} := (b, l_b, r_b)$. The estimated output is given by

$$\begin{aligned}(F_j, l_{F_j}, r_{F_j}) &= (a, l_a, r_a) \oplus (b, l_b, r_b) \ln(j+1)\\ &= (a + b\ln(j+1), l_a + l_b \ln(j+1), r_a + r_b \ln(j+1)).\end{aligned}$$

For the estimation of the parameters estimates of the mode and both the left and right spread are needed. According to Ishibuchi and Nii (2001) the modes are estimated with OLS in the same way as in Section 4.5.2 and we denote the estimated mode of the FN $\tilde{a}$ by $\hat{a}$ and the estimated mode of the parameter $\tilde{b}$ by $\hat{b}$. The spreads of the estimated parameters are determined with the help of the following linear program.

$$\min_{l_a, l_b, r_a, r_b} J l_a + l_b \sum_{j=0}^{J-1} \ln(j+1) + J r_a + r_b \sum_{j=0}^{J-1} \ln(j+1) \tag{4.27a}$$

s.t.

$$\begin{aligned}\hat{a} + \hat{b}\ln(j+1) - (l_a + l_b \ln(j+1))(1-\alpha^*) &\le \underline{F}_j\\ j &= 0, \dots, J-1\end{aligned} \tag{4.27b}$$

$$\hat{a} + \hat{b}\ln(j+1) + (r_a + r_b \ln(j+1))(1-\alpha^*) \geq \overline{F}_j \qquad j = 0,\dots,J-1 \tag{4.27c}$$

$$l_a, l_b, r_a, r_b \geq 0 \tag{4.27d}$$

After determining the parameters $\tilde{a}$ and $\tilde{b}$ they are converted back to development factors by

$$\tilde{f}_j := 1 + e^{\tilde{a}}(j+1)^{\tilde{b}}, \qquad j \in \{0,\dots,J-1\}. \tag{4.28}$$

Equation (4.28) not necessarily yields a TFN such that Andrés Sánchez proposes to make use of Taylor's expansion as described in Remark 2.27. The development factor $\tilde{f}_j$ can be approximated by a TFN $\tilde{f}' = (f_j, l_{f_j}, r_{f_j})$ considering the function $g : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto g(x,y) := 1 + e^x(j+1)^y$, where

$$\begin{aligned}
f_j &= 1 + e^{\hat{a}}(j+1)^{\hat{b}} \\
l_{f_j} &= \frac{\partial g(\hat{a},\hat{b})}{\partial x} l_a + \frac{\partial g(\hat{a},\hat{b})}{\partial y} l_b \\
&= e^{\hat{a}}(j+1)^{\hat{b}} l_a + e^{\hat{a}}(j+1)^{\hat{b}} \ln(j+1) l_b = e^{\hat{a}}(j+1)^{\hat{b}} \left(l_a + l_b \ln(j+1)\right) \\
r_{f_j} &= \frac{\partial g(\hat{a},\hat{b})}{\partial x} r_a + \frac{\partial g(\hat{a},\hat{b})}{\partial y} r_b \\
&= e^{\hat{a}}(j+1)^{\hat{b}} r_a + e^{\hat{a}}(j+1)^{\hat{b}} \ln(j+1) r_b = e^{\hat{a}}(j+1)^{\hat{b}} \left(r_a + r_b \ln(j+1)\right).
\end{aligned}$$

With the approximated development factors $\tilde{f}$ one can fill up the unobserved part of the claims development triangle by successively multiplying the last observation with the estimated TFN-development factors $\tilde{f}'$.[41] Furthermore, the predictions for the reserve can be deduced.

4.6.3 Fuzzy Taylor's Separation Method by Andrés Sánchez

The claims reserving method presented in Andrés Sánchez (2007) is a combination of the fuzzy regression method by Ishibuchi and Nii (2001) (cf. also the presentation on

[41] Andrés Sánchez (2006) makes an intermediate step by first calculating so-called "projection rates" $\tilde{f}_{j,s} = \prod_{h=j}^{s-1} \tilde{r}_h$, $j, s = 0,\dots,J-1$, $s > j$, to specifically predict certain values in the claims $\tilde{C}_{i,s}$, $i = 0,\dots,I$, in the development triangle.

pp. 53ff.) and Taylor's geometric separation method (see Section 4.5.3). A comparable approach in a shortened version is pursued in Bahrami and Bahrami (2015).

Taylor's method is extended in the way that the number of claims is assumed to be a TFN $\tilde{N}_i := (n_i, l_{n_i}, r_{n_i})$. By $Z_{i,j}$, $i, j \in \{0, \dots, J\}, i < j$, we denote paid claims made in accident year i and development year j. Then, the average cost per claim also is a TFN, i.e. $\tilde{S}_{i,j} := \frac{Z_{i,j}}{\tilde{N}_i}$. Since the procedure presented is based on Taylor's geometric separation method which is using logarithmized magnitudes, the logarithmized average cost per claim need to be considered. Buckley and Qu (1990) proposed to make use of α-cuts in order to derive a function of a TFN. Then, the average cost per claim can be represented as

$$\left(\ln \tilde{S}_{i,j}\right)_\alpha := \left[\underline{\ln S_{i,j}}(\alpha), \overline{\ln S_{i,j}}(\alpha)\right] := \left[\ln\left(\frac{Z_{i,j}}{n_i + r_{n_i}(1-\alpha)}\right), \ln\left(\frac{Z_{i,j}}{n_i - l_{n_i}(1-\alpha)}\right)\right]$$

(cf. Andrés Sánchez (2007, p. 152) and Buckley and Qu (1990, pp. 309f.)). In analogy to Taylor's model it is assumed that the average costs $\tilde{S}_{i,j}$ are determined by two factors

$$\tilde{S}_{i,j} = \tilde{P}_j \tilde{\Pi}_{i+j} \tag{4.29}$$

where $\tilde{P}_j$, $j \in \{0, \dots, J\}$, and $\tilde{\Pi}_{i+j}$, $i, j \in \{0, \dots, J\}, i < j$, are assumed to be FNs. Analogous to the crisp case, the factor $\tilde{\Pi}_{i+j}$ represents the inflation effect in calendar year $i + j$ and $\tilde{P}_j$ quantifies the share of claims being settled in development year j. A linear expression is yielded by applying logarithms to Equation (4.29), i.e.

$$\ln\left(\tilde{S}_{i,j}\right) = \ln\left(\tilde{P}_j\right) + \ln\left(\tilde{\Pi}_{i+j}\right).$$

The author assumes that both summands $\ln\left(\tilde{P}_j\right) := (\ln p_j, l_{\ln p_j}, r_{\ln p_j})$ and $\ln\left(\tilde{\Pi}_{i+j}\right) := (\ln \pi_{i+j}, l_{\ln \pi_{i+j}}, r_{\ln \pi_{i+j}})$ are TFNs. Consequently, $\ln\left(\tilde{S}_{i,j}\right)$ is a TFN as well as a sum of two TFNs. Therefore, we have

$$\begin{aligned}
\ln\left(\tilde{S}_{i,j}\right) &:= \left(\ln s_{i,j}, l_{\ln s_{i,j}}, r_{\ln s_{i,j}}\right) = \left(\ln p_j, l_{\ln p_j}, r_{\ln p_j}\right) \oplus \left(\ln \pi_{i+j}, l_{\ln \pi_{i+j}}, r_{\ln \pi_{i+j}}\right) \\
&= \left(\ln p_j + \ln \pi_{i+j}, l_{\ln p_j} + l_{\ln \pi_{i+j}}, r_{\ln p_j} + r_{\ln \pi_{i+j}}\right).
\end{aligned}$$

According to the FR technique proposed by Ishibuchi and Nii the modes of the coefficients are estimated with OLS and denoted by $\ln \hat{p}_j$ and $\ln \hat{\pi}_{i+j}$. The spreads are estimated by solving the following linear program given in Equations (4.30a)-(4.30d).

Hence, an estimator for the logarithmized average cost per claim $\ln(\tilde{S}_{i,j})$ is given by $\ln\left(\hat{\tilde{S}}_{i,j}\right) := \left(\ln \hat{s}_{i,j}, l_{\ln \hat{s}_{i,j}}, r_{\ln \hat{s}_{i,j}}\right)$. The factor $\ln \tilde{P}_0$ is assumed to be a crisp zero, i.e. $\ln \tilde{P}_0 := (0,0,0)$. Here, minimizing the total spreads means taking the sum over all observations in the claims development triangle (cf. Table 4.2).

accident-	development year j						
year i	0	1	$\cdots$	j	$\cdots$	$J-1$	J
0	$\ln \tilde{p}_0 + \ln \tilde{\pi}_0$	$\ln \tilde{p}_1 + \ln \tilde{\pi}_1$	$\cdots$	$\ln \tilde{p}_j + \ln \tilde{\pi}_j$	$\cdots$	$\ln \tilde{p}_{J-1} + \ln \tilde{\pi}_{J-1}$	$\ln \tilde{p}_J + \ln \tilde{\pi}_J$
1	$\ln \tilde{p}_0 + \ln \tilde{\pi}_1$	$\ln \tilde{p}_1 + \ln \tilde{\pi}_2$	$\cdots$	$\ln \tilde{p}_j + \ln \tilde{\pi}_{j+1}$	$\cdots$	$\ln \tilde{p}_{J-1} + \ln \tilde{\pi}_J$	
$\vdots$	$\vdots$				$\iddots$		
i	$\ln \tilde{p}_0 + \ln \tilde{\pi}_i$	$\ln \tilde{p}_1 + \ln \tilde{\pi}_{i+1}$	$\cdots$				
$\vdots$	$\vdots$		$\iddots$				
$I-1$	$\ln \tilde{p}_0 + \ln \tilde{\pi}_{I-1}$	$\ln \tilde{p}_1 + \ln \tilde{\pi}_J$					
I	$\ln \tilde{p}_0 + \ln \tilde{\pi}_J$						

Table 4.2: Claims development triangle with assumed logarithmized average claim costs in the model of Andrés Sánchez (2007).

There are $J+1-j$ terms $\ln p_j$ in each column j (starting from column one) and the term $\ln \pi_k$ occurs $k+1$-times in the k-th diagonal (counting the diagonals from the upper left corner starting with $k=0$). This leads to the following minimization problem:

$$\min_{l_{\ln p_j}, r_{\ln p_j}, l_{\ln \pi_k}, r_{\ln \pi_k}} \sum_{j=1}^{J} (J+1-j)\left(l_{\ln p_j} + r_{\ln p_j}\right) + \sum_{k=0}^{J} (k+1)\left(l_{\ln \pi_k} + r_{\ln \pi_k}\right) \quad (4.30a)$$

s.t.

$$\mathbf{A}_i \mathbf{v}_* \leq \underline{\ln S_i} \qquad i = 0,1,\ldots,I \quad (4.30b)$$

$$\mathbf{A}_i \mathbf{v}^* \geq \overline{\ln S_i} \qquad i = 0,1,\ldots,I \quad (4.30c)$$

$$l_{\ln p_j}, r_{\ln p_j}, l_{\ln \pi_k}, r_{\ln \pi_k} \geq 0 \qquad j = 1,\ldots,J, k = 0,1,\ldots,J \quad (4.30d)$$

where

$$\mathbf{A}_i := \left(\begin{array}{cc|c|c} \multicolumn{2}{c|}{\mathbf{0}_{1\times J}} & \mathbf{0}_{(J+1-i)\times i} & \mathbf{I}_{(J+1-i)\times(J+1-i)} \\ \mathbf{I}_{(J-i)\times(J-i)} & \mathbf{0}_{(J-i)\times i} & & \end{array}\right)$$

and

$$\mathbf{v}_* := \begin{pmatrix} \ln \hat{p}_1 - l_{\ln p_1}(1-\alpha*) \\ \ln \hat{p}_2 - l_{\ln p_2}(1-\alpha*) \\ \vdots \\ \ln \hat{p}_J - l_{\ln p_J}(1-\alpha*) \\ \ln \hat{\pi}_0 - l_{\ln \pi_0}(1-\alpha*) \\ \ln \hat{\pi}_1 - l_{\ln \pi_1}(1-\alpha*) \\ \vdots \\ \ln \hat{\pi}_J - l_{\ln \pi_J}(1-\alpha*) \end{pmatrix}, \qquad \mathbf{v}^* := \begin{pmatrix} \ln \hat{p}_1 + r_{\ln p_1}(1-\alpha*) \\ \ln \hat{p}_2 + r_{\ln p_2}(1-\alpha*) \\ \vdots \\ \ln \hat{p}_J + r_{\ln p_J}(1-\alpha*) \\ \ln \hat{\pi}_0 + r_{\ln \pi_0}(1-\alpha*) \\ \ln \hat{\pi}_1 + r_{\ln \pi_1}(1-\alpha*) \\ \vdots \\ \ln \hat{\pi}_J + r_{\ln \pi_J}(1-\alpha*) \end{pmatrix},$$

$$\underline{\ln S_i} := \begin{pmatrix} \underline{\ln S_{i,0}}(\alpha^*) \\ \underline{\ln S_{i,1}}(\alpha^*) \\ \vdots \\ \underline{\ln S_{i,J-i}}(\alpha^*) \end{pmatrix}, \qquad \overline{\ln S_i} := \begin{pmatrix} \overline{\ln S_{i,0}}(\alpha^*) \\ \overline{\ln S_{i,1}}(\alpha^*) \\ \vdots \\ \overline{\ln S_{i,J-i}}(\alpha^*) \end{pmatrix}$$

$\mathbf{I}_{i\times j}$ is an identity matrix with i rows and j columns, $\mathbf{0}_{i\times j} \in \mathbb{R}^{i\times j}$ is a matrix containing only zeros with i rows and j columns and $\alpha^* \in (0,1)$. The matrix $\mathbf{A}_0$ yields the conditions for row 0 in the claims development triangle, the matrix $\mathbf{A}_1$ for the first row and so on. Precisely speaking, the observed values should lie within the support of the estimated TFN, i.e. it is required that the difference of the estimated mode and the estimated left spread is less or equal than the α-cut of the observed average cost per claim and that the sum of the estimated mode and estimated right spread is greater or equal than the α-cut of the observed average cost per claim (cf. Equations (4.30b)-(4.30c)). Equation (4.30d) technically requires that both spreads are positive or zero.

With the above stated procedure estimators for the fuzzy parameters $\ln \tilde{p}_j$, $j \in \{1,\ldots,J\}$, and $\ln \tilde{\Pi}_k$, $k = 0,\ldots,J$ are yielded which belong to the model describing the upper left part of the triangle. In order to predict the ultimate claims estimators for $\ln \tilde{\Pi}_k$, $k = J+1,\ldots,2J$ are also needed. Again, Andrés Sánchez consults the FR technique by Ishibuchi and Nii (2001). It is assumed that the inflation factors can be represented in the following way

$$\ln \tilde{\Pi}_{i+j} = \tilde{A} \oplus \tilde{B}(i+j), \qquad i+j \leq J, \tag{4.31}$$

where $\ln \tilde{\Pi}_{i+j} := (\ln \Pi_{i+j}, l_{\ln \Pi_{i+j}}, r_{\ln \Pi_{i+j}})$, $\tilde{A} := (a, l_a, r_a)$ and $\tilde{B} := (b, l_b, r_b)$ are TFNs. Consequently, Equation (4.31) can be written as

$$\begin{aligned} \ln \tilde{\Pi}_{i+j} &= (\ln \Pi_{i+j}, l_{\ln \Pi_{i+j}}, r_{\ln \Pi_{i+j}}) = (a, l_a, r_a) \oplus (b, l_b, r_b)(i+j) \\ &= (a + b(i+j), l_a + l_b(i+j), r_a + r_b(i+j)). \end{aligned} \tag{4.32}$$

In this model this leads to a spread of $l_a + r_a + (i+j)(l_b + r_b)$ for one point in time. When fitting condition (4.31) to the observations $\ln \Pi_{i+j}$, $i + j = 0, \ldots, J$, the total spread of the output is (with $c := i + j$)

$$\begin{aligned} &\sum_{c=0}^{J} (l_a + r_a + c(l_b + r_b)) = (J+1)(l_a + r_a) + (l_b + r_b) \sum_{c=0}^{J} c \\ =&(J+1)(l_a + r_a) + (l_b + r_b)\frac{J(J+1)}{2} = (J+1)\left(l_a + r_a + (l_b + r_b)\frac{J}{2}\right). \end{aligned}$$

As a consequence the linear problem in order to determine the spreads is formulated according to Andrés Sánchez (2007, pp. 154f.) in the following way.[42] Within this the OLS estimates for the modes are denoted by $\hat{a}$ and $\hat{b}$.[43]

$$\min_{l_a, r_a, l_b, r_b} l_a + r_a + \frac{J}{2}(l_b + r_b) \tag{4.33a}$$

s.t.

$$\begin{aligned} \hat{a} - l_a + (\hat{b} - l_b)(i+j)(1-\alpha^*) &\leq \ln \hat{\Pi}_{i+j} - l_{\ln \Pi_{i+j}}(1-\alpha^*) \\ & i + j = 0, 1, \ldots, J \end{aligned} \tag{4.33b}$$

$$\begin{aligned} \hat{a} + r_a + (\hat{b} + r_b)(i+j)(1-\alpha^*) &\geq \ln \hat{\Pi}_{i+j} + r_{\ln \Pi_{i+j}}(1-\alpha^*) \\ & i + j = 0, 1, \ldots, J \end{aligned} \tag{4.33c}$$

$$l_a, r_a, l_b, r_b \geq 0 \tag{4.33d}$$

[42] The multiplicative constant can be left out in the minimization problem.

[43] Within the constraints of the linear problem it is surprising that the author does not multiply the spreads resulting form Equation (4.32) by the factor $(1-\alpha^*)$ but the estimator $\hat{b}$. It seems worth considering to replace constraints (4.33b) and (4.33c) by

$$\hat{a} + \hat{b}(i+j) - (l_a + l_b(i+j))\,(1-\alpha^*) \leq \ln \hat{\Pi}_{i+j} - l_{\ln \Pi_{i+j}}(1-\alpha^*), \quad i+j = 0, 1, \ldots, J$$

and

$$\hat{a} + \hat{b}(i+j) + (r_a + r_b(i+j))\,(1-\alpha^*) \geq \ln \hat{\Pi}_{i+j} + r_{\ln \Pi_{i+j}}(1-\alpha^*), \quad i+j = 0, 1, \ldots, J,$$

respectively.

Subsequent to determining the inflation factors the predicted logarithmized average claims payments $\ln \hat{\tilde{S}}_{i,j}$, $i \in \{1, \ldots, J\}$, $j \geq J - i + 1$, are derived by calculating

$$\ln \hat{\tilde{S}}_{i,j} = \left(\ln \hat{s}_{i,j}, l_{\ln \hat{s}_{i,j}}, r_{\ln \hat{s}_{i,j}}\right) = \left(\ln p_j + \ln \hat{\Pi}_{i+j}, l_{\ln p_j} + l_{\ln \hat{\Pi}_{i+j}}, r_{\ln p_j} + r_{\ln \hat{\Pi}_{i+j}}\right).$$

Retransformating these values leads to the predicted average claims payments

$$\hat{\tilde{S}}_{i,j} = \exp\left(\ln \hat{\tilde{S}}_{i,j}\right) = \exp\left(\ln\left(\tilde{P}_j\right) \oplus \ln\left(\hat{\tilde{\Pi}}_{i+j}\right)\right).$$

In general, the predicted values $\hat{\tilde{S}}_{i,j}$ are not TFNs but can be approximated by a TFN $\hat{\tilde{S}}_{i,j}^{\text{approx}}$ using the approximation presented in Remark 2.27 considering the function $g : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto \exp(x + y)$. With $\hat{\tilde{S}}_{i,j}^{\text{approx}} := (\hat{s}_{i,j}, l_{\hat{s}_{i,j}}, r_{\hat{s}_{i,j}})$ they get:

$$\begin{aligned}
\hat{s}_{i,j} &= \exp(\ln \hat{s}_{i,j}) = \exp(\ln p_j + \ln \hat{\Pi}_{i+j}) \\
l_{\hat{s}_{i,j}} &= \frac{\partial g(\ln \hat{s}_{i,j})}{\partial x} l_{\ln \hat{s}_{i,j}} = \exp(\ln \hat{s}_{i,j}) l_{\ln \hat{s}_{i,j}} = \exp(\ln p_j + \ln \hat{\Pi}_{i+j}) \left(l_{\ln p_j} + l_{\ln \hat{\Pi}_{i,j}}\right) \\
r_{\hat{s}_{i,j}} &= \frac{\partial g(\ln \hat{s}_{i,j})}{\partial y} r_{\ln \hat{s}_{i,j}} = \exp(\ln \hat{s}_{i,j}) r_{\ln \hat{s}_{i,j}} = \exp(\ln p_j + \ln \hat{\Pi}_{i+j}) \left(r_{\ln p_j} + r_{\ln \hat{\Pi}_{i,j}}\right)
\end{aligned}$$

This step is followed by calculating the predicted claims payments $\hat{\tilde{Z}}_{i,j}$ by using the relation $\hat{\tilde{Z}}_{i,j} \approx \hat{\tilde{S}}_{i,j} \otimes \tilde{N}_i$. This yields

$$\hat{\tilde{C}}_{i,j} \approx (\hat{s}_{i,j} n_i, \hat{s}_{i,j} l_{n_i} + n_i l_{\hat{s}_{i,j}}, \hat{s}_{i,j} r_{n_i} + n_i r_{\hat{s}_{i,j}}).$$

4.6.4 Application of Fuzzy Regression to Kremer's Two Way of ANOVA by Andrés Sánchez

A further application of the FR method by Ishibuchi and Nii (2001) is presented in Andrés Sánchez (2012). The procedure is based on a reserving technique proposed by Kremer (1982) (see Section 4.5.4).

In order to get a similar structure as in Kremer (1982) it is assumed that the incremental claims $\tilde{X}_{i,j}$ can be written as (cf. Equation (4.23))

$$\tilde{X}_{i,j} = \tilde{Z}_i \otimes \tilde{p}_j,$$

where again tildes indicate FNs. $\tilde{Z}_i$ stands for total claims payments in accident year i and $\tilde{p}_j$ denotes the ratio of claims payments done in development year j. Then, the following linear structure is assumed for the logarithmized incremental claims

$$\ln \tilde{X}_{i,j} = \tilde{a} \oplus \tilde{b}_i \oplus \tilde{c}_j \tag{4.34}$$

where $\tilde{a} := (a, l_a, r_a)$, $\tilde{b}_i := (b_i, l_{b_i}, r_{b_i})$ and $\tilde{c}_j := (c_j, l_{c_j}, r_{c_j})$ are TFNs. Consequently, $\ln \tilde{X}_{i,j}$ is a TFN as sum of TFNs as well and from Equation (4.34) they yield:

$$\begin{aligned}\ln \tilde{X}_{i,j} &= (\ln X_{i,j}, l_{\ln X_{i,j}}, r_{\ln X_{i,j}}) \\ &= (a, l_a, r_a) \oplus (b_i, l_{b_i}, r_{b_i}) \oplus (c_j, l_{c_j}, r_{c_j}) \\ &= (a + b_i + c_j, l_a + l_{b_i} + l_{c_j}, r_a + r_{b_i} + r_{c_j})\end{aligned}$$

Then, the modes are estimated by means of OLS whereas the spreads are yielded with the help of the following linear program (cf. Andrés Sánchez 2012, p. 2438). The estimates of the modes of the estimated parameters $\hat{\tilde{a}}$, $\hat{\tilde{b}}_i$ and $\hat{\tilde{c}}_j$ are denoted by $\hat{a}$, $\hat{b}_i$ and $\hat{c}_j$. Let $\alpha^* \in (0, 1)$.

$$\begin{aligned}\min_{l_a, l_{b_i}, l_{c_j}, r_a, r_{b_i}, r_{c_j}} & \frac{(J+1)^2 + (J+1)}{2} (l_a + r_a) + \sum_{i=1}^{I} (J + 1 - i) (l_{b_i} + r_{b_i}) \\ & + \sum_{j=1}^{J} (I + 1 - j) \left(l_{c_j} + r_{c_j}\right) \\ = & \frac{(J+2)(J+1)}{2} (l_a + r_a) + \sum_{i=1}^{I} (J + 1 - i) (l_{b_i} + r_{b_i}) \\ & + \sum_{j=1}^{J} (I + 1 - j)(l_{c_j} + r_{c_j})\end{aligned} \tag{4.35a}$$

s.t.

$$\hat{a} + \hat{b}_i + \hat{c}_j - \left(l_a + l_{b_i} + l_{c_j}\right) (1 - \alpha^*) \leq \ln X_{i,j}, \quad i = 1, \dots, I, j = 1, \dots, J - i \tag{4.35b}$$

$$\hat{a} + \hat{b}_i + \hat{c}_j + \left(r_a + r_{b_i} + r_{c_j}\right) (1 - \alpha^*) \geq \ln X_{i,j}, \quad i = 1, \dots, I, j = 1, \dots, J - i \tag{4.35c}$$

$$\hat{a} + \hat{b}_i + \hat{c}_j - \left(l_a + l_{b_i} + l_{c_j}\right) \geq 0, \quad i = 1, \dots, I, j = 1, \dots, J - i \tag{4.35d}$$

$$l_a, l_{b_i}, l_{c_j}, r_a, r_{b_i}, r_{c_j} \geq 0, \qquad i = 1, \dots, I, j = 1, \dots, J \tag{4.35e}$$

accident-	development year j						
year	0	1	$\cdots$	j	$\cdots$	$J-1$	J
0	$\tilde{a}$	$\tilde{a} \oplus \tilde{c}_1$	$\cdots$	$\tilde{a} \oplus \tilde{c}_j$	$\cdots$	$\tilde{a} \oplus \tilde{c}_{J-1}$	$\tilde{a} \oplus \tilde{c}_J$
1	$\tilde{a} \oplus \tilde{b}_1$	$\tilde{a} \oplus \tilde{b}_1 \oplus \tilde{c}_1$	$\cdots$	$\tilde{a} \oplus \tilde{b}_1 \oplus \tilde{c}_j$	$\cdots$	$\tilde{a} \oplus \tilde{b}_1 \oplus \tilde{c}_{J-1}$	
$\vdots$	$\vdots$				$\iddots$		
i	$\tilde{a} \oplus \tilde{b}_i$	$\tilde{a} \oplus \tilde{b}_i \oplus \tilde{c}_1$	$\cdots$				
$\vdots$	$\vdots$						
			$\iddots$				
$I-1$	$\tilde{a} \oplus \tilde{b}_{J-1}$	$\tilde{a} \oplus \tilde{b}_{I-1} \oplus \tilde{c}_1$					
I	$\tilde{a} \oplus \tilde{b}_I$						

Table 4.3: Logarithmized incremental claims cost $\ln \tilde{X}_{i,j}$ in a development triangle under the assumptions of the model by Andrés Sánchez (2012).

In Equation (4.35a) the total spreads (over the run-off triangle) are minimized. The linear factor $\tilde{a}$ occurs $\frac{(J+2)(J+1)}{2}$ times, the linear factor $\tilde{b}_i$, $i = 1, \ldots, I$ appears $(J+1-i)$ times in every row beginning from row one and the linear factor $\tilde{c}_j$ turns up $(I+1-j)$ times in every column beginning from column one (see Table 4.3). Conditions (4.35b) and (4.35c) ensure that the estimated values (adjusted by the h-certain factor) comprise the observed values. Equation (4.35d) requires that the sum of the linear factors is non-negative and Equation (4.35e) yields that the linear factors are non-negative.

As an extension to the fuzzy ANOVA model presented above Andrés Sánchez (2014) propose a first attempt to model inflation in fuzzy claims reserving methods.[44] The interest rate is assumed to be a TFN $\tilde{\rho}$ given by $\tilde{\rho} := (\rho, l_\rho, r_\rho)$. Andrés Sánchez points to the Institute of Actuaries and assumes that claims payments are done at mid-year. Then the present value of a predicted reserve is calculated by

$$\tilde{s}_{i,j} \otimes e^{\tilde{\rho}\left(J+\frac{1}{2}-(i+j)\right)}.$$

Subsequently, the present value is defuzzified with the help of the concept of the expected value of a FN in order to determine a reserve.

[44] We do not dedicate a separate section to the presentation of the work by Andrés Sánchez (2014) since to our understanding the only new aspect in comparison to Andrés Sánchez (2012) is the extension to the incorporation of interest rates.

4.6.5 Hybrid Fuzzy Least-Squares Regression Analysis for the Geometric Separation Method

In their 2010 work Başer and Apaydin claim that there might be a loss of reliability of statistical methods in "the presence of real world factors" (cf. Başer and Apaydin 2010, p. 113). The authors make use of the fuzzy regression method as proposed by Chang (2001) whose goal is to consider both randomness and fuzziness and apply it to the geometric separation method by Taylor (see Section 4.5.3). Furthermore, the confidence intervals of the parameter estimates in a regression context are utilized to construct TFNs.

As a first step, average losses $S_{i,j} := \frac{Z_{i,j}}{N_i}$ for every observation in the upper left part of the triangle consisting of paid losses are calculated where N_i denotes the number of claims in accident year i and $Z_{i,j}$ denote claims payments in accident year i and development year j.[45] According to the geometric separation method a multiplicative structure is assumed, i.e.

$$S_{i,j} = r_j \lambda_{i+j}, , \qquad \forall i, j, i + j \leq J, \tag{4.36}$$

where λ_{i+j} is a inflation factor for calendar year $i + j$ and the factor r_j describes the share of claims being settled in development year j. To get a linear additive structure logarithms are applied to Equation (4.36), i.e.

$$\ln S_{i,j} = \ln r_j + \ln \lambda_{i+j}, \qquad \forall i, j, i + j \leq J. \tag{4.37}$$

An error term $\ln \varepsilon_{i,j}$ is added to the right side of Equation (4.37) in order to apply OLS.

$$\ln S_{i,j} = \ln r_j + \ln \lambda_{i+j} + \ln \varepsilon_{i,j}, \qquad \forall i, j, i + j \leq J, \tag{4.38}$$

where the error terms $\ln \varepsilon_{i,j}$ are assumed to be i.i.d. and following a normal distribution. Assuming the linear structure as in Equation (4.38) estimations $\widehat{\ln r_j}$ and $\widehat{\ln \lambda_{i-j}}$ of the parameters $\ln r_j$ and $\ln \lambda_{i+j}$ are derived with OLS. For better readability they are in the following denoted by $\ln \hat{r}_j$ and $\ln \hat{\lambda}_{i+j}$. Subsequently, the coefficients of

[45] Başer and Apaydin (2010) write that they consider incurred losses. However, we assume that paid losses are being considered.

the regression model are fuzzified by using the confidence intervals of the parameter estimators deduced before. According to the multiple linear regression model the t-statistics follows a t-distribution with $\frac{1}{2}(J^2-J-2)$ degrees of freedom[46], i.e. this leads to

$$\frac{\ln \hat{r}_j - \ln r_j}{\mathrm{se}(\ln \hat{r}_j)} \sim t\left(\tfrac{1}{2}(J^2-J-2)\right) \quad \text{and} \quad \frac{\ln \hat{\lambda}_{i+j} - \ln \lambda_{i+j}}{\mathrm{se}(\ln \hat{\lambda}_{i+j})} \sim t\left(\tfrac{1}{2}(J^2-J-2)\right).$$

Consequently, a $(1-\gamma)$-confidence interval can be constructed for the regression coefficients $\ln \hat{r}_j$ and $\ln \hat{\lambda}_{i+j}$ for a fix $\gamma \in (0,1)$. These are given by

$$\left[\ln \hat{r}_j - t_{\frac{1}{2}(J^2-J-2),1-\gamma/2}\mathrm{se}\left(\ln \hat{r}_j\right), \ln \hat{r}_j + t_{\frac{1}{2}(J^2-J-2),1-\gamma/2}\mathrm{se}\left(\ln \hat{r}_j\right)\right]$$

and

$$\left[\ln \hat{\lambda}_{i+j} - t_{\frac{1}{2}(J^2-J-2),1-\gamma/2}\mathrm{se}\left(\ln \hat{\lambda}_{i+j}\right), \ln \hat{\lambda}_{i+j} + t_{\frac{1}{2}(J^2-J-2),1-\gamma/2}\mathrm{se}\left(\ln \hat{\lambda}_{i+j}\right)\right].$$

With the help of the confidence intervals STFNs $\ln \hat{\tilde{r}}_j$ and $\ln \hat{\tilde{\lambda}}_{i+j}$ are constructed:[47] the centers of the symmetrical confidence intervals mark the modes and half of the width of the support specifies the left and right spread, respectively. Consequently, we yield

$$\ln \hat{\tilde{r}}_j = \left(\ln \hat{r}_j, l_{\ln \hat{r}_j}\right) := \left(\ln \hat{r}_j, t_{(J^2-J-2)\frac{1}{2},1-\frac{\gamma}{2}}\mathrm{se}(\ln \hat{r}_j)\right)$$

and

$$\ln \hat{\tilde{\lambda}}_{i+j} = \left(\ln \hat{\lambda}_{i+j}, l_{\ln \hat{\lambda}_{i+j}}\right) := \left(\ln \hat{\lambda}_{i+j}, t_{(J^2-J-2)\frac{1}{2},1-\frac{\gamma}{2}}\mathrm{se}(\ln \hat{\lambda}_{i+j})\right).$$

As a means to predict future incurred claims a similar approach as in Andrés Sánchez (2007) (cf. Section 4.6.3) is applied. Hybrid fuzzy least squares regression (cf. p. 57ff.) is used to fit a fuzzy regression model with symmetrical coefficients to the logarithmized inflation parameters. Hence, we consider

[46] Başer and Apaydin claim that the t-statistics are follow a t-distribution with $J-2$ degrees of freedom (cf. Başer and Apaydin 2010, p. 118). However, we adapted the model in the way that the degrees of freedom are changed. In Taylor's geometric separation method we have $\frac{1}{2}(J+1)(J+2)$ observations in the triangle and $(2J+1)$ unknown parameters. Thus, the t-statistic is t-distributed with $\frac{1}{2}(J+1)(J+2)-(2J+1)-1=\frac{1}{2}(J^2-J-2)$ degrees of freedom.

[47] This follows the idea of Buckley (2006) where TFNs are being built from confidence intervals for a binomially distributed random variable (cf. Buckley 2006, 21f.). Moreover, fuzzy coefficients are constructed from confidence intervals of the regression coefficients in a simple linear regression model (cf. Buckley 2006, 171f.).

$$\ln \hat{\tilde{\lambda}}_{i+j} = \tilde{a} \oplus \tilde{b}(i+j), i+j \leq J,$$

where $\tilde{a} := (a, l_a)$ and $\tilde{b} := (b, l_b)$ are STFNs. Therefore, the predicted values can be represented as

$$\begin{aligned}\ln \hat{\tilde{\lambda}}_{i+j} &= \left(\ln \hat{\lambda}_{i+j}, l_{\ln \hat{\lambda}_{i+j}}\right) = (a, l_a) \oplus (b, l_b)(i+j) \\ &= (a + b(i+j), l_a + l_b(i+j)), \qquad i+j = J+1, \dots, 2J.\end{aligned} \tag{4.39}$$

According to the hybrid fuzzy least squares method (cf. Equations (2.17) and (2.18)) the modes and spreads of the fuzzy coefficients are determined with the help of the two following linear systems of Equations (4.40) and (4.41).

$$\begin{array}{ccccc} a(J+1) & + & b\sum_{i+j=0}^{J}(i+j) & = & \sum_{i+j=0}^{J} \ln \hat{\lambda}_{i+j} \\ a\sum_{i+j=0}^{J}(i+j) & + & b\sum_{i+j=0}^{J}(i+j)^2 & = & \sum_{i+j=0}^{J}(i+j) \ln \hat{\lambda}_{i+j} \end{array} \tag{4.40}$$

$$\begin{array}{ccccc} (J+1)l_a & + & l_b\sum_{i+j=0}^{J}(i+j) & = & \sum_{i+j=0}^{J} l_{\ln \hat{\lambda}_{i+j}} \\ l_a\sum_{i+j=0}^{J}(i+j) & + & l_b\sum_{i+j=0}^{J}(i+j)^2 & = & \sum_{i+j=0}^{J}(i+j) l_{\ln \hat{\lambda}_{i+j}} \end{array} \tag{4.41}$$

Using Equation (4.39) the estimated parameters $\ln \hat{\tilde{\lambda}}_{i+j}$, $i+j = J+1, \dots, 2J$, are obtained. Applying the estimated coefficients $\ln \hat{\tilde{r}}_j$, $j = 1, \dots, J$, and $\ln \hat{\tilde{\lambda}}_{i+j}$, $i+j = J+1, \dots, 2J$, the lower right part of the triangle is predicted, i.e.

$$\ln \hat{\tilde{S}}_{i,j} = \left(\ln \hat{S}_{i,j}, l_{\ln \hat{S}_{i,j}}\right) = \ln \hat{\tilde{r}}_j \oplus \ln \hat{\tilde{\lambda}}_{i+j} = \left(\ln \hat{r}_j + \ln \hat{\lambda}_{i+j}, l_{\ln \hat{r}_j} + l_{\ln \hat{\lambda}_{i+j}}\right)$$

The average claims are obtained by retransformation of the logarithmized average claims, i.e.

$$\hat{\tilde{S}}_{i,j} = \exp\left(\ln \hat{\tilde{S}}_{i,j}\right).$$

In order to do so Başer and Apaydin make use of a defuzzification method following weighted fuzzy arithmetic as introduced by Chang (2001) (cf. Başer and Apaydin 2010, p. 116). By this means, crisp predictions of the average claims in the lower right part of the development triangle are yielded (cf. Başer and Apaydin 2010, p. 119)

$$\hat{S}_{i,j} = \exp\left(\ln \hat{S}_{i,j}\right)\left(\frac{\exp(l_{\ln \hat{S}_{i,j}}) + \exp(-l_{\ln \hat{S}_{i,j}}) - 2}{l^2_{\ln \hat{S}_{i,j}}}\right).$$

Utilizing the relation $\hat{Z}_{i,j} = \hat{S}_{i,j} N_i$ predictions for the outstanding claims payments are obtained. Then, the reserves for single and aggregated accident years are derived.

4.6.6 Further Methods

In addition to the methods presented before some more models and fuzzy approaches in claims reserving techniques have been published. The aim of this section is to give a short overview of these methods.

The publications by Kerkez (2013a,b) join earlier articles in the way that fuzzy regression is presented as a means to predict the reserves if information is vague or insufficient. Both methods apply the fuzzy regression model by Tanaka et al. (1982) with symmetric coefficients.[48]

Kerkez assumes that the average development of the claims from one development year to the next can be described with the help of a linear model, i.e.

$$C_{i,j+1} = a_j + b_j C_{i,j} + \varepsilon_{i,j},$$

where $a_j, b_j \in \mathbb{R}$ and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, $\sigma^2 > 0$. Then, the model is transferred to a fuzzy one by choosing fuzzy symmetrical coefficients $\tilde{a}_j := (a_j, l_{a_j})$ and $\tilde{b}_j := (b_j, l_{b_j})$. In this way, the fuzzy linear model is given by

$$\tilde{C}_{i,j+1} = \tilde{a}_j + \tilde{b}_j \tilde{C}_{i,j}, \qquad i = 0, \ldots, I - j - 1.$$

The estimated modes and the spreads of the fuzzy coefficient and fuzzy intercept are determined with the model of Tanaka et al. (1982) (cf. Equations (2.12a)-(2.12d)). The estimated values are referred to as $\hat{\tilde{a}}_j$ and $\hat{\tilde{b}}_j$. With these notations a prediction of the ultimate claim can de deduced with the following recursive formula:

[48] Both publications – to our state of view – do not substantially enhance the state of literature and, therefore, are only briefly mentioned here. However, they do address a related problem to that studied in Chapter 6. That is the reason why we include it in the overview.

$$\hat{\hat{C}}_{i,J} = \hat{\hat{a}}_{J-1} \oplus \hat{\hat{b}}_{J-1} \left(\ldots \hat{\hat{a}}_{j+1} \oplus \hat{\hat{b}}_{j+1} \left(\hat{\hat{a}}_j \oplus \hat{\hat{b}}_j C_{i,j} \right) \right) \tag{4.42}$$

The prediction of the ultimate claim in Equation (4.42) does not necessarily lead to a STFN. Therefore, the approximation by Dubois is used to derive a STFN (cf. Remark 2.27).

To allow for a comparison between the classical CL-model and a fuzzy counterpart Kerkez fits a regression model without and with intercept to a given data set in a numerical example.[49] The proposed procedure finishes with a deduction of a fuzzy reserve.

Onoghojobi and Olewuezi (2015) aim to use a fuzzy regression model combined with the CL reserving method. Closer examination shows that the authors indeed calculate reserves based on the CL method. Moreover, they do a fuzzy regression on a given data set of settled claims. However, a fuzzy regression in combination with the CL method is not investigated.[50]

The publication of Kim and Kim (2014) introduces a method using a fuzzy version of the Hoerl curve. The Hoerl curve is introduced e.g. in Kaas et al. (2008, p. 276). They extend the method of Andrés Sánchez (2006) (see also Section 4.6.2) by also taking into account a calendar year effect.
In order to calculate development year factors a fuzzy log Hoerl curve is considered, i.e. the incremental claims are fit to the following model:

$$\tilde{Z}_{i,j} = \tilde{c} + \tilde{\alpha}_i + \tilde{\beta} \log(j) + \tilde{\gamma}_j,$$

where $\tilde{c}, \tilde{\alpha}_i, \tilde{\beta}, \tilde{\gamma}_j$ are FNs. In order to deduce the fuzzy parameters a fuzzy version of the analysis of covariance is presented (cf. Kim and Kim 2014, pp. 352f.). It makes use of OLS as well as fuzzy regression techniques. Taking into account the approximations of the incremental claims with the help of the Hoerl curve the development factors are deduced. In the following, fuzzy reserves are determined.

[49] A more thorough explanation of a fuzzy version of the CL-model is given in Section 6.2.
[50] The article lacks some explanation on the chosen data and interpretation. E.g. the CL-reserves are calculated on a calender year basis and compared with reserves for an occurence year (cf. Onoghojobi and Olewuezi 2015, p. 21). Moreover, the transition from a fuzzy regression model with asymmetric spreads to one with symmetric spreads is not comprehensible in our opinion (cf. Onoghojobi and Olewuezi 2015, p. 22).

4.6.7 Discussion

To our best knowledge there are still few applications of fuzzy methods in the context of claims reserving. This is surprising since subjective judgment and intuition is also present in this field. Further reasons for the incorporation of fuzzy methods in this field have been mentioned in Chapter 1. Comparing the applications in claims reserving leads to the observation that most applications so far are utilizations of FR (see Andrés Sánchez and Terceño Gómez (2003), Andrés Sánchez (2006, 2007, 2012, 2014), Başer and Apaydin (2010), and Kerkez (2013a,b) as presented in Sections 4.6.1-4.6.6). However, some stochastic reserving methods employ regression techniques (see e.g. Taylor (1978) and Benjamin and Eagles (1986)) but a large part within the set of stochastic reserving methods does not.[51] According to the present status of knowledge the use of fuzzy methods in claims reserving resorts to FR:

The method proposed by Andrés Sánchez and Terceño Gómez has been the first approach in 2003 even though there are good reasons which allow for the use of fuzzy methodologies in claims reserving. They present a simple model making use of the FR technique as presented in Tanaka and Ishibuchi (1992). Consequently, the fuzzy regression parameters are symmetric such that large deviations in only one direction can not be captured by the model. This may lead to large total spreads of the regression parameters and, thus, to an overestimation of the reserves. Moreover, the proposed method asks for a specification of the h-certain factor with the choice of the factor h not being elaborated on.[52] Furthermore, the reserves are only specified as TFNs, even though "fuzzy" reserves can not be set up by an actuary. Additionally, the uncertainty of the prediction is not quantified.

The subsequent publication by Andrés Sánchez (2006) makes use of FR as suggested by Ishibuchi and Nii (2001) to avoid the disadvantages of Tanaka's 1987 model. The use of asymmetric TFNs circumvents the drawbacks of symmetric parameters. The author brings up as a motivation that the cardinality of the observed data set might be too small to reasonably conduct OLS estimates. However, he chooses the method

[51] Applications of fuzzy methods not using FR will be discussed in Chapters 5 and 7.
[52] The choice of the h-certain factor is discussed in Moskowitz and Kim (1993).

by Ishibuchi and Nii which uses OLS as an estimation method for the mode of the parameters.

The same applies for the article by Andrés Sánchez (2007). Additionally, the procedure presented here can be divided into two parts: First, fuzzy parameters are adjusted with the help of FR regression and the given development triangle. As a second step the inflation parameters are adjusted to a fuzzy regression model. In particular, not all information is used as the authors employs only the modes of the fuzzy inflation factors. Consequently, the information about the spreads gets lost at this point. However, it is the first approach to include inflation effects in the field of fuzzy claims reserving.

Andrés Sánchez (2012) (as well as the extension presented in Andrés Sánchez (2014)) employ similar techniques as the publications discussed before and also fall back on the fuzzy regression model by Ishibuchi and Nii (2001). Başer and Apaydin (2010) follow new paths with an approach to consider both randomness and fuzziness as their model makes use of the hybrid fuzzy least squares regression. By choosing the confidence interval the variability and, hence, the fuzziness, can be chosen by the user as they are used to construct TFNs. Moreover, they apply a defuzzification method by Chang and Ayyub (2001).

For the fuzzy reserving methods by Kerkez (2013a,b) the before discussed drawbacks of symmetrical coefficients occur by means of the usage of the fuzzy regression method by Tanaka and Ishibuchi (1992). Nevertheless, it represents an approach to consider the most popular reserving method, i.e. the CL method, within the “fuzzy” world which has been not regarded before (except for the special case in the London Chain). Onoghojobi and Olewuezi (2015) aim to use fuzzy regression methods in a CL context. However, closer examination shows that it is not conducted.

The concentration on regression approaches is a major difference to other existing claims reserving methods among which a large part is not based on regression approaches. In the later chapters 5 and 7 this gap will be closed with the presentation of two regression-free fuzzy claims reserving methods.

5 | The Fuzzy Chain-Ladder Model – An Approach with Fuzzy Numbers

This chapter is based on a joint work with Jochen Heberle of which an earlier version is published in *Insurance: Mathematics & Economics* (cf. Heberle and Thomas 2014).[53]

Firstly, in Section 5.1 motives for making use of fuzzy set theory are given and the basic notation is introduced. In the subsequent Section 5.2 the fuzzy chain-ladder model is presented and basic results are shown. Section 5.3 then deals with the resulting reserves in our proposed method and the uncertainty of the ultimate claims is derived in Section 5.4. In Section 5.5 an example is presented. Lastly, a discussion of the presented method is given in Section 5.6.

5.1 Motivation and Notation

As pointed out in Chapter 4 non-life actuaries need to predict the claims reserves in an appropriate way. There exists a variety of models in claims reserving which are either purely computational or of stochastic nature and make use of so-called development factors which aim to model the average increase from one development year to another. Both classes of claims reserving methods yield crisp magnitudes. However, non-life actuaries might tend to adjust development factors due to their subjective judgment.

[53]Parts of this paper were presented at the ASTIN Colloquium in The Hague, The Netherlands, 2013, at the IFSA World Congress/NAFIPS Annual meeting in Edmonton, Canada, 2013 and the APRIA Conference in New York, USA, 2013.

In the case of the popular chain-ladder method (cf. Section 4.3) this means that the CL-factors are modified subsequently. More precisely, if the calculated value of a development factor is c, the adapted value can be perceived as approximately c. In this way, vagueness is included in the considered magnitudes which can not be captured by ordinary stochastic methods. Hence, we here propose a fuzzy chain-ladder model in which no subsequent alteration of the CL-factors is needed, and, nevertheless, the development factors are flexible and include uncertainty.

The use of fuzzy set theory for applications in insurance is not new but started off in the 1980's. A detailed overview is given in Chapter 3 where it can be seen that the field of loss reserving has not received much attention. In contrast to the publications as elaborated on in Chapter 3 our model is not an application of fuzzy regression but makes use of fuzzy numbers and its corresponding arithmetic.[54] Furthermore, we make an attempt to quantify the prediction uncertainty and not only give a prediction for the ultimate claim.

The results in the fuzzy chain-ladder model are yielded by using triangular fuzzy numbers which can be classified as a special case of L-R FNs. With the help of FNs no information about the uncertainty of a CL-factor is lost. In this chapter one possibility to estimate the development factors is presented and we conduct the analysis with this choice. As we use TFNs for modeling the CL-factors the mode as well as the left and right spread need to be estimated.

Even though there exists a number of claims reserving methods the CL method is the most widely used method in practice due to its simplicity and nonetheless good results. As in the previous chapters we denote by $C_{i,j}$ cumulative claims made in relative accident year $i \in \{0, \dots, I\}$ and development year $j \in \{0, \dots, J\}$. At time $t = I$ (calendar year) we have the following set of observations (cf. Equation (4.1)):

$$\mathcal{D}_I = \{C_{i,j} \mid i + j \leq I\}$$

The set of observations is given in the upper left part of a development triangle as depicted in Figure 4.2. As the CL-factors f_j are not observable in practice they are estimated in the following way (cf. Equation (4.5)):

[54] An application of fuzzy regression methods to the CL procedure is presented in Chapter 6.

$$\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}, \quad j = 0, \ldots, J-1$$

The core idea in this chapter is to model the CL-factors f_j which are fraught with vagueness due to subjective judgment by fuzzy numbers. For simplicity's sake triangular fuzzy numbers (cf. Definition 2.18) are taken since they are both easy to handle arithmetically and to interpret. As fuzzy multiplication the secant approximation as deduced in Equation (2.7) is applied. Moreover, we make use of the defuzzification methods, i.e. the expected value of a TFN (cf. Definition 2.38), and uncertainty measures, i.e. the uncertainty of a FN (cf. Definition 2.50), as described before.

5.2 The Model

Our objective is to develop a predictor for the ultimate claims $C_{i,J}$, $i = 1, \ldots, I$, using the classical chain-ladder method in combination with fuzzy numbers. In this way we derive a model which copes with the vagueness in the CL-factors due to subjective judgment. We use the notations as introduced in Section 5.1. Furthermore we derive in Section 5.4 an estimator for the uncertainty of the predicted ultimate claims on the basis of Model Assumptions 5.1.

In the case of adjustment of the CL-factors vagueness is added to them. Hence, we propose TFNs to model the development factors. In this situation the width of the left and right spread represents the vagueness.

Model Assumptions 5.1 (Fuzzy chain-ladder (FCL) model)

- There exist triangular fuzzy numbers $\tilde{f}_j$ $(j = 0, \ldots, J-1)$ so that the cumulative claims can be expressed as

  $$\tilde{C}_{i,j+1} = \tilde{f}_j \tilde{C}_{i,j}$$

  for all $i = 0, \ldots, I$ and $j = 0, \ldots, J-1$.

- The sums of incremental claims $\bigoplus_{i=0}^{I-j-1} \tilde{X}_{i,j+1}$ with $\tilde{X}_{i,j+1} = \tilde{C}_{i,j+1} \ominus \tilde{C}_{i,j}$ for all $j \in \{0, \ldots, J-1\}$ are non-negative.[55]

[55] In the original publication Heberle and Thomas (2014) it is assumed that incremental claims $\tilde{X}_{i,j+1}$ are non-negative. However, this assumption can be relaxed as stated in Model Assumptions 5.1.

Remarks 5.2

- According to Equation (2.2) we can perceive a real number as a fuzzy number. In the special case of TFN a real number $a \in \mathbb{R}$ can be denoted as $\tilde{a} := (a, 0, 0)$. Thus, we can write all observable cumulative claims $\tilde{C}_{i,j}$, i.e. if $i + j \leq I$ holds, as $\tilde{C}_{i,j} = (C_{i,j}, 0, 0)$.

- In comparison to the Model Assumptions for the classical CL method e.g. given in Merz and Wüthrich (2007) our model does not have an error term ε_j.[56] This is due to the fact that in the FCL model the uncertainty, i.e. the vagueness, is completely contained in the fuzzy numbers. As pointed out in Section 2.1.4, in a fuzzy environment vagueness is modelled but no randomness in a stochastic sense.

With Model Assumptions 5.1 we can achieve the following result regarding the expected value and uncertainty of a cumulative claim $\tilde{C}_{i,j+1}$ given an observation $\tilde{C}_{i,j}$. We make use of the notations and definitions as introduced in Definition 2.38, Remark 2.39a) and b) as well as Definition 2.50.

Lemma 5.3

For all $0 \leq i \leq I$ and $0 \leq j \leq J - 1$ let $\tilde{C}_{i,j}$ and $\tilde{f}_j = (f_j, l_{f_j}, r_{f_j})$ be positive fuzzy numbers. Let $\beta \in [0, 1]$ and $K \in \mathbb{R}^+$. For $i + j \leq I$ the cumulative claims $\tilde{C}_{i,j}$ are observable and, hence, crisp, i.e. $\tilde{C}_{i,j} = (C_{i,j}, 0, 0)$. Then, we have for $i, j, i + j \leq I$,

$$E_\beta[\tilde{C}_{i,j+1} \mid \tilde{C}_{i,j}] = C_{i,j} \left(f_j - \frac{1-\beta}{2} l_{f_j} + \frac{\beta}{2} r_{f_j} \right) \tag{5.1}$$

and

$$\mathrm{Unc}_K(\tilde{C}_{i,j+1} \mid \tilde{C}_{i,j}) = \tfrac{1}{2} K C_{i,j} (l_{f_j} + r_{f_j}). \tag{5.2}$$

Proof. We have

$$\begin{aligned} E_\beta[\tilde{C}_{i,j+1} \mid \tilde{C}_{i,j}] &= E_\beta[\tilde{f}_j \tilde{C}_{i,j} \mid \tilde{C}_{i,j}] \\ &= E_\beta[(f_j, l_{f_j}, r_{f_j}) \otimes (C_{i,j}, 0, 0)] \\ &= E_\beta[(f_j C_{i,j}, l_{f_j} C_{i,j}, r_{f_j} C_{i,j})] \end{aligned}$$

[56] The same model is also considered in e.g. Gisler and Wüthrich (2008) and Peters et al. (2010).

$$= f_j C_{i,j} - \frac{1-\beta}{2} l_{f_j} C_{i,j} + \frac{\beta}{2} r_{f_j} C_{i,j}$$

and

$$\begin{aligned}
\mathrm{Unc}_K(\tilde{C}_{i,j+1} \mid \tilde{C}_{i,j}) &= \mathrm{Unc}_K(\tilde{f}_j \tilde{C}_{i,j} \mid \tilde{C}_{i,j}) \\
&= \mathrm{Unc}_K((f_j, l_{f_j}, r_{f_j})(C_{i,j}, 0, 0)) \\
&= \mathrm{Unc}_K((f_j C_{i,j}, l_{f_j} C_{i,j}, r_{f_j} C_{i,j})) \\
&= \tfrac{1}{2} K C_{i,j} (l_{f_j} + r_{f_j}).
\end{aligned}$$

□

Remark 5.4
The model assumptions in the classical chain-ladder model are (cf. Model Assumptions 4.2 and Wüthrich and Merz (2008, p. 37)):

$$\begin{aligned}
E[C_{i,j+1} \mid C_{i,0}, \dots, C_{i,j}] &= f_j C_{i,j} = \; E[C_{i,j+1} \mid C_{i,j}] \\
\mathrm{Var}(C_{i,j+1} \mid C_{i,0}, \dots, C_{i,j}) &= \sigma_j^2 C_{i,j} = \mathrm{Var}(C_{i,j+1} \mid C_{i,j})
\end{aligned} \tag{5.3}$$

For $\beta = 0.5$ and equal spreads, i.e. $l_{f_j} = r_{f_j}$, Equation (5.1) results in $E_\beta[\tilde{C}_{i,j+1} \mid \tilde{C}_{i,j}] = f_j C_{i,j}$. In Equation (5.2) $\frac{1}{2} K(l_{f_j} + r_{f_j})$ plays a similar role as σ_j^2 in (5.3). In the previous case it is a measure for uncertainty whereas in the latter case it is the variance parameter. Hence, the defined expected value and uncertainty in this model deal with uncertainty e.g. due to subjective judgment while the variance parameter considers randomness as in classical statistics / probability theory (cf. Section 2.1.4).

In the following an estimator for the fuzzy chain-ladder factors $\tilde{f}_j$, $j = 0, \dots, J-1$, is proposed. As these factors are TFNs, estimators for the mode, left and right spread are needed. Throughout this chapter estimators are denoted with a "hat".

An actuary will choose the width of the spreads according to the confidence he/she has in the given data. The more reliable the data is, the narrower he/she will choose the spreads due to higher confidence. In the same way, if he/she has less trust in the data, he/she will assess broader spreads. The spreads do not necessarily be equal but can be of different length. However, a nice result is yielded when choosing the estimator as in Estimator 5.5. The definition of the estimator is based on the following consideration:

The estimator of the mode is given by the classical one as it might be adjusted. If there is no profound knowledge available about the data, an actuary will prefer wider spreads. Model Assumptions 5.1 lead to a limiting case as the sum of incremental claims is assumed to be non-negative. In the following Definition 5.5 the left spreads cannot go further down as one since

$$\frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} - \frac{\sum_{i=0}^{I-j-1} X_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} = \frac{\sum_{i=0}^{I-j-1}(C_{i,j+1} - C_{i,j+1} + C_{i,j})}{\sum_{i=0}^{I-j-1} C_{i,j}} = 1.$$

For simplicity reasons an symmetrical approach is used.

Estimator 5.5 (FCL estimator for the TFNs $\tilde{f}_j$)
The fuzzy chain-ladder factors $\tilde{f}_j = (f_j, l_{f_j}, r_{f_j})$, $j = 0, \ldots, J-1$, as introduced in Model Assumptions 5.1 are estimated by $\hat{\tilde{f}}_j = (\hat{f}_j, \hat{l}_{\hat{f}_j}, \hat{r}_{\hat{f}_j})$ with

$$\begin{aligned} \hat{f}_j &:= \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \\ \hat{l}_{\hat{f}_j} := \hat{r}_{\hat{f}_j} &:= \frac{\sum_{i=0}^{I-j-1} X_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \end{aligned} \tag{5.4}$$

where $X_{i,j}$ are incremental claims for $i = 0, \ldots, I$ and $j = 0, \ldots, J-1$.

Remarks 5.6

- The fuzzy chain-ladder $\tilde{f}_j$ factor can attain a value of one and a little above with low positive grade of membership. The mode $\hat{f}_j$ assumes a degree of membership of one.

- The second postulation in Model Assumptions 5.1 assures that the spreads of the estimators are non-negative because the numerator in (5.4) is non-negative.

- Since we are choosing $\hat{l}_{\hat{f}_j} = \hat{r}_{\hat{f}_j}$ for all $j = 0, \ldots, J-1$ the estimator yields symmetric TFNs. This definition is not mandatory but it leads to simpler results. By doing so, we assume that up-/downward variations in f_j are equally likely.

- The total spread of $\hat{\tilde{f}}_j$ given by $\hat{l}_{\hat{f}_j} + \hat{r}_{\hat{f}_j}$ can be considered as the "uncertainty" of the FN $\hat{\tilde{f}}_j$. The broader the whole spread $\hat{l}_{\hat{f}_j} + \hat{r}_{\hat{f}_j}$, the more uncertain the factor $\hat{\tilde{f}}_j$.

- The modes and expected values $E_\beta(\hat{\tilde{f}}_j)$, $j = 0, \dots, J-1$, are equal to the classical chain-ladder estimators (cf. Wüthrich and Merz 2008, p. 38).

As an analogue to the predictor in the classical case we are capable of predicting the ultimate claims $\tilde{C}_{i,J}$ $(i = 1, \dots, I)$ by

$$\widehat{\tilde{C}}_{i,J} := \tilde{C}_{i,I-i} \otimes \bigotimes_{j=I-i}^{J-1} \hat{\tilde{f}}_j. \tag{5.5}$$

In the above predictor the cumulative claims $C_{i,I-i}$, $i = 1, \dots, I$ are observable. Hence, they are crisp.

5.3 Claims Reserves

In Equation (5.5) a predictor for the ultimate claim is given. With the help of that definition we can derive a predictor for outstanding loss liabilities, i.e. a predictor for the claims reserve for single accident years (cf. Equation (4.2) for the crisp case). The fuzzy claims reserve $\tilde{R}_i$ for a fixed accident year $i = 1, \dots, I$ can be determined by

$$\tilde{R}_i = \tilde{C}_{i,J} \ominus \tilde{C}_{i,I-i}$$

where $\tilde{C}_{i,I-i}$ is observable, i.e. $\tilde{C}_{i,I-i} = (C_{i,I-i}, 0, 0)$. At time $t = I$ the ultimate claims $\tilde{C}_{i,J}$ are not observable and therefore, need to be predicted. With Equation (5.5) a predictor for the reserve $\tilde{R}_i$ is given by

$$\hat{\tilde{R}}_i = \hat{\tilde{C}}_{i,J} \ominus \tilde{C}_{i,I-i} = \tilde{C}_{i,I-i} \left(\bigotimes_{j=I-i}^{J-1} \hat{\tilde{f}}_j - 1 \right)$$

for $i = 1, \dots, I$.

The aggregated outstanding loss liabilities for all accident years are given by

$$\tilde{R} = \bigoplus_{i=1}^{I} \tilde{R}_i = \bigoplus_{i=1}^{I} \left(\tilde{C}_{i,J} - \tilde{C}_{i,I-i} \right)$$

and we yield a predictor for the aggregated outstanding loss liabilities by

$$\hat{\tilde{R}} = \bigoplus_{i=1}^{I} \hat{\tilde{R}}_i = \bigoplus_{i=1}^{I} \left(\tilde{C}_{i,I-i} \left(\bigotimes_{j=I-i}^{J-1} \hat{\tilde{f}}_j - 1 \right) \right).$$

Since actuaries cannot set up a fuzzy value in the balance sheet one needs to make use of defuzzification methods as e.g. the expected value of a fuzzy number (cf. Definition 2.38) for a given decision maker risk parameter.

5.4 Prediction Uncertainty

In this section we quantify the prediction uncertainty of the ultimate claim as given in Equation (5.5). In Section 5.4.1 single accident years are considered whereas in Section 5.4.2 aggregated accident years are regarded.

5.4.1 Single Accident Years

In the classical chain-ladder model the conditional mean square error of prediction (MSEP) is a popular means for quantifying the prediction uncertainty (cf. Definition 4.1).

As shown in Wüthrich and Merz (2008, p. 45) the MSEP can be computed in different ways using conditional and unconditional resampling methods. Since the fuzzy chain-ladder framework differs from the classical one the calculation is done in a different way.

Due to the model set-up the last observation for a given accident year $i \in \{1, \ldots, I\}$ is multiplied with at least one TFN $\tilde{f}_j$, $j = 0, \ldots, J-1$, to yield the ultimate claim. Considering the multiplication of fuzzy numbers the vagueness increases for the product as the spreads (cf. Equation (2.7) for the secant approximation). Hence, the uncertainty for increasing accident years rises as well, because more fuzzy chain-ladder factors need to be multiplied.

Thus, a representation for the product of several fuzzy chain-ladder factor estimators $\hat{\tilde{f}}_j$, $j = 0, \ldots, J-1$, is needed prior to deriving an estimator for the prediction uncertainty.

Lemma 5.7 (Fuzzy product of several chain-ladder factor estimators)
Let $\hat{\tilde{f}}_k$, $k = 0, \ldots, J-1$, *be estimators of the fuzzy chain-ladder factors* $\tilde{f}_k$, $k = 0, \ldots, J-1$. *Then, we have*

$$\bigotimes_{k=I-i}^{j} \hat{\tilde{f}}_k = \left(\hat{F}^j_{I-i}, \hat{l}_{\hat{F}^j_{I-i}}, \hat{r}_{\hat{F}^j_{I-i}}\right), \qquad j = I-i, \ldots, J-1,$$

with

$$\begin{aligned} \hat{F}^j_{I-i} &:= \prod_{k=I-i}^{j} \hat{f}_k \\ \hat{l}_{\hat{F}^j_{I-i}} &:= \hat{F}^j_{I-i} - 1 \\ \hat{r}_{\hat{F}^j_{I-i}} &:= \prod_{k=I-i}^{j} \left(2\hat{f}_k - 1\right) - \hat{F}^j_{I-i} \end{aligned}$$

for $i = 1, \ldots, I$.

Proof. Let $\tilde{A} = (a, l_a, r_a)$ and $\tilde{B} = (b, l_b, r_b)$ be two fuzzy numbers. Another representation $\tilde{A}'$ for $\tilde{A}$ is given if the entries in $\tilde{A}$ do not denote the mode, left and right spread but the position on the x-axis, i.e.

$$\tilde{A}' := (a - l_a, a, a + r_a)' =: (d, e, f)'.$$

An illustration is given in Figure 5.1. In the same way we write $\tilde{B}$ as $\tilde{B}' = (b - l_b, b, b + r_b)' =: (g, h, i)'$. The equivalent notation is always denoted with an apostrophe. With

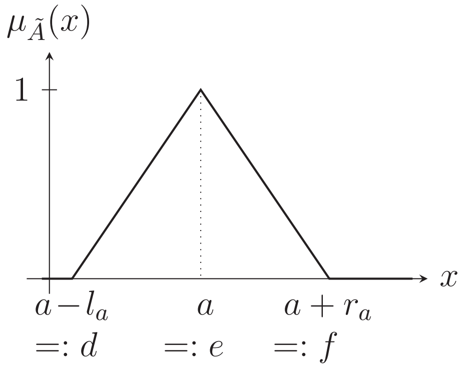


Figure 5.1: Equivalent notation for a TFN $\tilde{A} = (a, l_a, r_a)$.

the approximated multiplication as defined in Equation (2.7) (i.e. secant approximation) we yield in the equivalent representation

$$\begin{aligned}\tilde{A}' \otimes \tilde{B}' &= (ab - (al_b + bl_a - l_a l_b), ab, ab + (ar_b + br_a + r_a r_b))' \\ &= (gd, eh, if)'.\end{aligned}$$

Every estimator $\hat{\hat{f}}_j$ $(j = 0, \ldots, J-1)$ can be represented as (cf. Estimator 5.5):

$$\hat{\hat{f}}_j = (\hat{f}_j, \hat{f}_j - 1, \hat{f}_j - 1) = (1, \hat{f}_j, 2\hat{f}_j - 1)'$$

Therefore, we get:

$$\begin{aligned}\bigotimes_{k=I-i}^{j} \hat{\hat{f}}_k &= \bigotimes_{k=I-i}^{j} \left(1, \hat{f}_k, 2\hat{f}_k - 1\right)' \\ &= \left(1, \prod_{k=I-i}^{j} \hat{f}_k, \prod_{k=I-i}^{j} \left(2\hat{f}_k - 1\right)\right)' \\ &= \left(\prod_{k=I-i}^{j} \hat{f}_k, \prod_{k=I-i}^{j} \hat{f}_k - 1, \prod_{k=I-i}^{j} \left(2\hat{f}_k - 1\right) - \prod_{k=I-i}^{j} \hat{f}_k\right)\end{aligned}$$

□

Remark 5.8

An iterative representation of the right spread $\hat{r}_{\hat{F}^{j+1}_{I-i}}$ is given by the recursion

$$\hat{r}_{\hat{F}^{j+1}_{I-i}} = \left(\hat{f}_{j+1} - 1\right) \hat{F}^{j}_{I-i} + \hat{f}_{j+1} r_{\hat{F}^{j}_{I-i}} + r_{\hat{F}^{j}_{I-i}} \left(\hat{f}_{j+1} - 1\right) \tag{5.6}$$

since

$$\begin{aligned}&\left(\hat{F}^{j}_{I-i}, \hat{l}_{\hat{F}^{j}_{I-i}}, \hat{r}_{\hat{F}^{j}_{I-i}}\right) \otimes \left(\hat{f}_{j+1}, \hat{f}_{j+1} - 1, \hat{f}_{j+1} - 1\right) \\ = &\left(\hat{F}^{j+1}_{I-i}, \hat{F}^{j+1}_{I-i} - 1, \hat{F}^{j}_{I-i}\left(\hat{f}_{j+1} - 1\right) + \hat{f}_{j+1} \hat{r}_{\hat{F}^{j}_{I-i}} + \hat{r}_{\hat{F}^{j}_{I-i}} \left(\hat{f}_{j+1} - 1\right)\right).\end{aligned}$$

Depending on the focus of consideration either the representation in Lemma 5.7, the following Corollary 5.9 or the recursion given in Equation (5.6) should be used.

Corollary 5.9

Another representation of the product of FCL factors as in Lemma 5.7 is given by

$$\left(\hat{f}_{I-i}, \hat{f}_{I-i} - 1, \hat{f}_{I-i} - 1\right) \otimes \ldots \otimes \left(\hat{f}_j, \hat{f}_j - 1, \hat{f}_j - 1\right)$$

$$
\begin{aligned}
=\Bigg(&\prod_{k=I-i}^{j}\hat{f}_k,\prod_{k=I-i}^{j}\hat{f}_k-1,\ \left(2^{j-(I-i)+1}-1\right)\prod_{k=I-i}^{j}\hat{f}_k\\
&+(-1)^{j-(I-i)+1}\sum_{k=1}^{j-(I-i)}(-2)^k\sum_{\substack{A\subseteq\{I-i,\ldots,j\}\\|A|=k}}\prod_{s\in A}\hat{f}_s-(-1)^{j-(I-i)}\Bigg).
\end{aligned}
$$

Proof. The proof is conducted via complete induction.

Base case:

Multiplication with one factor results in

$$
\begin{aligned}
&\left(\hat{f}_{I-i},\hat{f}_{I-i}-1,\hat{f}_{I-i}-1\right)\\
=&\left(\prod_{k=I-i}^{I-i}\hat{f}_k,\prod_{k=I-i}^{I-i}\hat{f}_k-1,(2-1)\hat{f}_{I-i}+0-1\right)\\
=&\Bigg(\prod_{k=I-i}^{I-i}\hat{f}_k,\prod_{k=I-i}^{I-i}\hat{f}_k-1,\ \left(2^{I-i-(I-i)+1}-1\right)\prod_{k=I-i}^{I-i}\hat{f}_k\\
&+(-1)^{I-i-(I-i)+1}\sum_{k=1}^{I-i-(I-i)}(-2)^k\sum_{\substack{A\subseteq\{I-i\}\\|A|=k}}\prod_{s\in A}\hat{f}_s-(-1)^{I-i-(I-i)}\Bigg).
\end{aligned}
$$

Induction hypothesis (IH):

The proposition holds true for multiplication with $j-(I-i)+1$ factors, i.e.

$$
\begin{aligned}
&\left(\hat{f}_{I-i},\hat{f}_{I-i}-1,\hat{f}_{I-i}-1\right)\otimes\ldots\otimes\left(\hat{f}_j,\hat{f}_j-1,\hat{f}_j-1\right)\\
=\Bigg(&\prod_{k=I-i}^{j}\hat{f}_k,\prod_{k=I-i}^{j}\hat{f}_k-1,\ \left(2^{j-(I-i)+1}-1\right)\prod_{k=I-i}^{j}\hat{f}_k\\
&+(-1)^{j-(I-i)+1}\sum_{k=1}^{j-(I-i)}(-2)^k\sum_{\substack{A\subseteq\{I-i,\ldots,j\}\\|A|=k}}\prod_{s\in A}\hat{f}_s-(-1)^{j-(I-i)}\Bigg).
\end{aligned}
$$

Inductive step:

We show that the proposition also holds for $j+1-(I-i)+1$ factors.

$$
\begin{aligned}
&\bigotimes_{k=I-i}^{j+1}\left(\hat{f}_k,\hat{f}_k-1,\hat{f}_k-1\right)\\
=&\bigotimes_{k=I-i}^{j}\left(\hat{f}_k,\hat{f}_k-1,\hat{f}_k-1\right)\otimes\left(\hat{f}_{j+1},\hat{f}_{j+1}-1,\hat{f}_{j+1}-1\right)
\end{aligned}
$$

$$\begin{aligned}
&\overset{\text{(IH)}}{=} \Bigg(\prod_{k=I-i}^{j} \hat{f}_k, \prod_{k=I-i}^{j} \hat{f}_k - 1, \left(2^{j-(I-i)+1} - 1\right) \prod_{k=I-i}^{j} \hat{f}_k \\
&\quad + (-1)^{j-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^k \sum_{\substack{A \subseteq \{I-i,\dots,j\} \\ |A|=k}} \prod_{s \in A} \hat{f}_s - (-1)^{j-(I-i)} \Bigg) \\
&\quad \otimes \left(\hat{f}_{j+1}, \hat{f}_{j+1} - 1, \hat{f}_{j+1} - 1\right)
\end{aligned}$$

We have a look at the centre, left and right spread of the resulting fuzzy number separately. For the centre we get

$$\prod_{k=I-i}^{j} \hat{f}_k \cdot \hat{f}_{j+1} = \prod_{k=I-i}^{j+1} \hat{f}_k.$$

The left spread results in

$$\begin{aligned}
&\prod_{k=I-i}^{j} \hat{f}_k (\hat{f}_{j+1} - 1) + \hat{f}_{j+1} \left(\prod_{k=I-i}^{j} \hat{f}_k - 1 \right) - \left(\prod_{k=I-i}^{j} \hat{f}_k - 1 \right) (\hat{f}_{j+1} - 1) \\
&= \prod_{k=I-i}^{j+1} \hat{f}_k - 1.
\end{aligned}$$

For the right spread we yield:

$$\begin{aligned}
&\prod_{k=I-i}^{j} \hat{f}_k (\hat{f}_{j+1} - 1) + \hat{f}_{j+1} \Bigg(\left(2^{j-(I-i)+1} - 1\right) \prod_{k=I-i}^{j} \hat{f}_k \\
&\quad + (-1)^{j-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^k \sum_{\substack{A \subseteq \{I-i,\dots,j\} \\ |A|=k}} \prod_{s \in A} \hat{f}_s - (-1)^{j-(I-i)} \Bigg) \\
&\quad + \Bigg(\left(2^{j-(I-i)+1} - 1\right) \prod_{k=I-i}^{j} \hat{f}_k \\
&\quad + (-1)^{j-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^k \sum_{\substack{A \subseteq \{I-i,\dots,j\} \\ |A|=k}} \prod_{s \in A} \hat{f}_s - (-1)^{j-(I-i)} \Bigg) (\hat{f}_{j+1} - 1) \\
&= \left(2 \cdot 2^{j-(I-i)+1} - 1\right) \prod_{k=I-i}^{j+1} \hat{f}_k + 2\hat{f}_{j+1} (-1)^{j-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^k \sum_{\substack{A \subseteq \{I-i,\dots,j\} \\ |A|=k}} \prod_{s \in A} \hat{f}_s \\
&\quad - (-1)^{j-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^k \sum_{\substack{A \subseteq \{I-i,\dots,j\} \\ |A|=k}} \prod_{s \in A} \hat{f}_s
\end{aligned}$$

$$
\begin{aligned}
&- 2^{j-(I-i)+1} \prod_{k=I-i}^{j} \hat{f}_k - 2\hat{f}_{j+1}(-1)^{j-(I-i)} - (-1)^{j+1-(I-i)} \\
&= \left(2^{j+1-(I-i)+1} - 1\right) \prod_{k=I-i}^{j+1} \hat{f}_k + (-1)^{j+1-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^{k+1} \sum_{\substack{A\subseteq\{I-i,\dots,j\} \\ |A|=k}} \hat{f}_{j+1} \prod_{s\in A} \hat{f}_s \\
&\quad + (-1)^{j+1-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^k \sum_{\substack{A\subseteq\{I-i,\dots,j\} \\ |A|=k}} \prod_{s\in A} \hat{f}_s \\
&\quad - 2^{j+1-(I-i)} \prod_{k=I-i}^{j} \hat{f}_k + (-1)^{j+1-(I-i)+1}(-2)\hat{f}_{j+1} - (-1)^{j+1-(I-i)} \\
&= \left(2^{j+1-(I-i)+1} - 1\right) \prod_{k=I-i}^{j+1} \hat{f}_k + (-1)^{j+1-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^{k+1} \sum_{\substack{A\subseteq\{I-i,\dots,j\}\times\{j+1\} \\ |A|=k+1}} \prod_{s\in A} \hat{f}_s \\
&\quad + (-1)^{j+1-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^k \sum_{\substack{A\subseteq\{I-i,\dots,j\} \\ |A|=k}} \prod_{s\in A} \hat{f}_s \\
&\quad - 2^{j+1-(I-i)} \prod_{k=I-i}^{j} \hat{f}_k + (-1)^{j+1-(I-i)+1}(-2)\hat{f}_{j+1} - (-1)^{j+1-(I-i)} \\
&= \left(2^{j+1-(I-i)+1} - 1\right) \prod_{k=I-i}^{j+1} \hat{f}_k + (-1)^{j+1-(I-i)+1} \sum_{k=2}^{j+1-(I-i)} (-2)^{k} \sum_{\substack{A\subseteq\{I-i,\dots,j\}\times\{j+1\} \\ |A|=k}} \prod_{s\in A} \hat{f}_s \\
&\quad + (-1)^{j+1-(I-i)+1} \sum_{k=1}^{j-(I-i)} (-2)^k \sum_{\substack{A\subseteq\{I-i,\dots,j\} \\ |A|=k}} \prod_{s\in A} \hat{f}_s \qquad (5.7) \\
&\quad + (-1)^{j+1-(I-i)+1}(-2)^{j+1-(I-i)} \prod_{k=I-i}^{j} \hat{f}_k + (-1)^{j+1-(I-i)+1}(-2)\hat{f}_{j+1} - (-1)^{j+1-(I-i)}
\end{aligned}
$$

$$
\begin{aligned}
&= \left(2^{j+1-(I-i)+1} - 1\right) \prod_{k=I-i}^{j+1} \hat{f}_k \\
&\quad + (-1)^{j+1-(I-i)+1} \sum_{k=1}^{j+1-(I-i)} (-2)^k \sum_{\substack{A\subseteq\{I-i,\dots,j+1\} \\ |A|=k}} \prod_{s\in A} \hat{f}_s - (-1)^{j+1-(I-i)} \qquad (5.8)
\end{aligned}
$$

The third line in (5.7) holds true since

$$
(-1)^{j+1-(I-i)+1}(-2)^{j+1-(I-i)} = (-1)^{j+1-(I-i)+1}(-1)^{j+1-(I-i)}2^{j+1-(I-i)}
$$

$$= (-1)^{2(j+1)-2(I-i)+1} 2^{j+1-(I-i)}$$
$$= -2^{j+1-(I-i)}.$$

In (5.7) the first sum comprises all k-fold products where $2 \le k \le j+1-(I-i)$ and $(j+1)$ is always a factor. The second sum contains all k-fold products where $1 \le k \le j-(I-i)$ but the factors are elements of $\{I-i,\dots,j\}$. Together with the product $\prod_{k=I-i}^{j} \hat{f}_k$ and the 1-fold product of the factor $\hat{f}_{j+1}$ (and its respective factors) this results in the sum in (5.8). □

With the help of Lemma 5.7 a predictor for the uncertainty of the predicted ultimate claims given the observations can be deduced. We make use of the notation in Definition 2.50.

Predictor 5.10 (Uncertainty of the ultimate claim)
Given the observations $\mathcal{D}_I$ and a parameter $K \in \mathbb{R}^+$ the uncertainty of the prediction of ultimate claims $\hat{\tilde{C}}_{i,J}$ for each accident year $i = 1,\dots,I$ is given by

$$\mathrm{Unc}_K\left(\hat{\tilde{C}}_{i,J} \mid \mathcal{D}_I\right) = \tfrac{1}{2} K \tilde{C}_{i,I-i}\left(\hat{l}_{\hat{F}^{J-1}_{I-i}} + \hat{r}_{\hat{F}^{J-1}_{I-i}}\right)$$

where $\tilde{C}_{i,I-i} := (C_{i,I-i}, 0, 0)$.

Proof. Let $K \in \mathbb{R}^+$. Given the observations $\mathcal{D}_I$ the predictor for the ultimate claims is known to be (cf. Equation (5.5))

$$\hat{\tilde{C}}_{i,J} = \tilde{C}_{i,I-i} \bigotimes_{j=I-i}^{J-1} \hat{\tilde{f}}_j.$$

At time $t = I$ all cumulative claims $\tilde{C}_{i,I-i}$ are known and observable. Therefore, they can be perceived as crisp numbers. Hence, the uncertainty of the predicted ultimate claims is given by:

$$\begin{aligned}
\mathrm{Unc}_K\left(\hat{\tilde{C}}_{i,J} \mid \mathcal{D}_i\right) &= \mathrm{Unc}_K\left(\tilde{C}_{i,I-i} \bigotimes_{j=I-i}^{J-1} \hat{\tilde{f}}_j \mid \mathcal{D}_i\right) \\
&= \mathrm{Unc}_K\left((C_{i,I-i}, 0, 0) \otimes \left(\hat{F}^{J-1}_{I-i}, \hat{l}_{\hat{F}^{J-1}_{I-i}}, \hat{r}_{\hat{F}^{J-1}_{I-i}}\right)\right) \\
&= \mathrm{Unc}_K\left(\left(C_{i,I-i}\hat{F}^{J-1}_{I-i}, C_{i,I-i}\hat{l}_{\hat{F}^{J-1}_{I-i}}, C_{i,I-i}\hat{r}_{\hat{F}^{J-1}_{I-i}}\right)\right)
\end{aligned}$$

$$= \tfrac{1}{2} K C_{i,I-i} \left(\hat{l}_{\hat{F}_{I-i}^{J-1}} + \hat{r}_{\hat{F}_{I-i}^{J-1}} \right)$$

□

The above-mentioned predictor takes into account single accident years while in Section 5.4.2 aggregated accident years are considered.

5.4.2 Aggregated Accident Years

The uncertainty of the predictor for aggregated ultimate claims $\bigoplus_{i=1}^{I} \hat{\tilde{C}}_{i,J}$ given the observations $\mathcal{D}_I$ is derived in a different way as in the classical CL-method since we are not dealing with random variables. The methods used here are equal to those in Section 5.4.1.

As a start we compute the sum of the ultimate claims over all accident years given the observations $\mathcal{D}_I$. As the observations in the claims development triangle are already known, we consider these as crisp numbers in the following Lemma 5.11.

Lemma 5.11
Given the observations $\mathcal{D}_I$ the sum of the predicted ultimate claims is given by

$$\bigoplus_{i=1}^{I} \hat{\tilde{C}}_{i,J} = \left(\sum_{i=1}^{I} C_{i,I-i} \hat{F}_{I-i}^{J-1}, \sum_{i=1}^{I} C_{i,I-i} \hat{l}_{\hat{F}_{I-i}^{J-1}}, \sum_{i=1}^{I} C_{i,I-i} \hat{r}_{\hat{F}_{I-i}^{J-1}} \right).$$

Proof. The sum of the predicted ultimate claims, given the observations $\mathcal{D}_I$, can be derived as:

$$\begin{aligned}
\bigoplus_{i=1}^{I} \hat{\tilde{C}}_{i,J} &= \bigoplus_{i=1}^{I} \left(\tilde{C}_{i,I-i} \bigotimes_{j=I-i}^{J-1} \hat{\tilde{f}}_j \right) \\
&= \bigoplus_{i=1}^{I} (C_{i,I-i}, 0, 0) \otimes \left(\hat{F}_{I-i}^{J-1}, \hat{l}_{\hat{F}_{I-i}^{J-1}}, \hat{r}_{\hat{F}_{I-i}^{J-1}} \right) \\
&= \bigoplus_{i=1}^{I} \left(C_{i,I-i} \hat{F}_{I-i}^{J-1}, C_{i,I-i} \hat{l}_{\hat{F}_{I-i}^{J-1}}, C_{i,I-i} \hat{r}_{\hat{F}_{I-i}^{J-1}} \right) \\
&= \left(\sum_{i=1}^{I} C_{i,I-i} \hat{F}_{I-i}^{J-1}, \sum_{i=1}^{I} C_{i,I-i} \hat{l}_{\hat{F}_{I-i}^{J-1}}, \sum_{i=1}^{I} C_{i,I-i} \hat{r}_{\hat{F}_{I-i}^{J-1}} \right)
\end{aligned}$$

With the help of Lemma 5.11 the uncertainty of the sum of the ultimate claims can be derived.

Estimator 5.12 (Uncertainty of the aggregated ultimate claims)
Given the observations $\mathcal{D}_I$ and a parameter $K \in \mathbb{R}^+$, the uncertainty of the aggregated ultimate claims $\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J}$ is given by:

$$\mathrm{Unc}_K\left(\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J} \mid \mathcal{D}_I\right) = \sum_{i=1}^{I} \mathrm{Unc}_K\left(\hat{\hat{C}}_{i,J} \mid \mathcal{D}_I\right)$$

Proof. With the notations as given in Lemma 5.7 we yield:

$$\begin{aligned}
\mathrm{Unc}_K\left(\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J} \mid \mathcal{D}_I\right) &= \mathrm{Unc}_K\left(\sum_{i=1}^{I} C_{i,I-i}\hat{F}_{I-i}^{J-1}, \sum_{i=1}^{I} C_{i,I-i}\hat{l}_{\hat{F}_{I-i}^{J-1}}, \sum_{i=1}^{I} C_{i,I-i}\hat{r}_{\hat{F}_{I-i}^{J-1}} \mid \mathcal{D}_I\right) \\
&= \frac{1}{2}K\left(\sum_{i=1}^{I} C_{i,I-i}\hat{l}_{\hat{F}_{I-i}^{J-1}} + \sum_{i=1}^{I} C_{i,I-i}\hat{r}_{\hat{F}_{I-i}^{J-1}}\right) \\
&= \frac{1}{2}K\Big(C_{1,I-1}\hat{l}_{\hat{F}_{I-1}^{J-1}} + C_{1,I-1}\hat{r}_{\hat{F}_{I-1}^{J-1}} + \dots \\
&\qquad + C_{I,0}\hat{l}_{\hat{F}_{0}^{J-1}} + C_{I,0}\hat{r}_{\hat{F}_{0}^{J-1}}\Big) \\
&= \sum_{i=1}^{I} \mathrm{Unc}_K\Big(\hat{\hat{C}}_{i,J} \mid \mathcal{D}_I\Big)
\end{aligned}$$

□

Thus, the uncertainty of ultimate claims for aggregated accidents is the sum of the uncertainty of ultimate claims for single accident years.

5.5 Example

In the following we apply the FCL method and investigate the run-off triangle given in Taylor and Ashe (1983). The triangle is presented in Table 5.1.
For this specific example we compute the fuzzy chain-ladder factors as given in Table 5.2.

The fuzzy development factors are written down in three lines. The first one indicates the mode which is in fact the classical CL-factor due to the chosen estimator. The

accident-year i	development year j 0	1	2	3	4	5	6	7	8	9
0	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
1	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	
2	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315		
3	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268			
4	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311				
5	396,132	1,333,217	2,180,715	2,985,752	3,691,712					
6	440,832	1,288,463	2,419,861	3,483,130						
7	359,480	1,421,128	2,864,498							
8	376,686	1,363,294								
9	344,014									

Table 5.1: Observed cumulative claims $C_{i,j}$.

second and third line show the left and right spread, respectively. Since the estimator in this method is chosen as a symmetrical triangular fuzzy number they are equal.

$\hat{\hat{f}}_j$	development year 0	1	2	3	4	5	6	7	8
$\hat{f}_j$	3.4906	1.7473	1.4574	1.1739	1.1038	1.0863	1.0539	1.0766	1.0177
$\hat{l}_{\hat{f}_j}$	2.4906	0.7473	0.4574	0.1739	0.1038	0.0863	0.0539	0.0766	0.0177
$\hat{r}_{\hat{f}_j}$	2.4906	0.7473	0.4574	0.1739	0.1038	0.0863	0.0539	0.0766	0.0177

Table 5.2: Fuzzy chain-ladder factors $\hat{\hat{f}}_j$ for $j = 0, \ldots, J-1$.

With the help of the computed fuzzy development factors we are able to predict the unobservable right part of the run-off triangle as shown in Table 5.3. The grey shaded part of the table indicates the observations which are crisp magnitudes. For each accident year i and development year j we write down the predicted cumulated claims in three lines. The first stands for the mode, and the second and third ones for the left and right spread of the cumulative figures, respectively. The first line for each accident year is equal to the results in the classical CL method. As the fuzzy multiplication is not a closed operation we use an approximation which again yields triangular fuzzy numbers that are not necessarily symmetrical. Hence, the values in the second and third line in Table 5.3 differ. In fact, the right spread is increasing more rapidly.

The estimated reserves for this specific example are derived with the methods as presented in Section 5.3 and given in Table 5.4. As the estimator for the FCL method is chosen according to Estimator 5.5 the left spread of the fuzzy development factor reaches to one (cf. Table 5.2). Therefore, the reserve in the FCL method can amount to

$\hat{\tilde{C}}_{i,j}$	development year j 0	1	2	3	4	5	6	7	8	9
$\hat{C}_{0,j}$	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
$\hat{l}_{\hat{C}_{0,j}}$	0	0	0	0	0	0	0	0	0	0
$\hat{r}_{\hat{C}_{0,j}}$	0	0	0	0	0	0	0	0	0	0
$\hat{C}_{1,j}$	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	5,433,719
$\hat{l}_{\hat{C}_{1,j}}$	0	0	0	0	0	0	0	0	0	94,634
$\hat{r}_{\hat{C}_{1,j}}$	0	0	0	0	0	0	0	0	0	94,634
$\hat{C}_{2,j}$	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315	5,285,148	5,378,826
$\hat{l}_{\hat{C}_{2,j}}$	0	0	0	0	0	0	0	0	375,833	469,511
$\hat{r}_{\hat{C}_{2,j}}$	0	0	0	0	0	0	0	0	375,833	482,834
$\hat{C}_{3,j}$	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268	4,835,458	5,205,637	5,297,906
$\hat{l}_{\hat{C}_{3,j}}$	0	0	0	0	0	0	0	247,190	617,369	709,638
$\hat{r}_{\hat{C}_{3,j}}$	0	0	0	0	0	0	0	247,190	655,217	770,712
$\hat{C}_{4,j}$	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311	4,207,459	4,434,133	4,773,589	4,858,200
$\hat{l}_{\hat{C}_{4,j}}$	0	0	0	0	0	0	334,148	560,822	900,278	984,889
$\hat{r}_{\hat{C}_{4,j}}$	0	0	0	0	0	0	334,148	596,826	1,027,662	1,148,703
$\hat{C}_{5,j}$	396,132	1,333,217	2,180,715	2,985,752	3,691,712	4,074,999	4,426,546	4,665,023	5,022,155	5,111,171
$\hat{l}_{\hat{C}_{5,j}}$	0	0	0	0	0	383,287	734,834	973,311	1,330,443	1,419,459
$\hat{r}_{\hat{C}_{5,j}}$	0	0	0	0	0	383,287	800,966	1,125,746	1,655,241	1,802,935
$\hat{C}_{6,j}$	440,832	1,288,463	2,419,861	3,483,130	4,088,678	4,513,179	4,902,528	5,166,649	5,562,182	5,660,771
$\hat{l}_{\hat{C}_{6,j}}$	0	0	0	0	605,548	1,030,049	1,419,398	1,683,519	2,079,052	2,177,641
$\hat{r}_{\hat{C}_{6,j}}$	0	0	0	0	605,548	1,155,789	1,744,557	2,196,651	2,928,515	3,130,917
$\hat{C}_{7,j}$	359,480	1,421,128	2,864,498	4,174,756	4,900,545	5,409,337	5,875,997	6,192,562	6,666,635	6,784,799
$\hat{l}_{\hat{C}_{7,j}}$	0	0	0	1,310,258	2,036,047	2,544,839	3,011,499	3,328,064	3,802,137	3,920,301
$\hat{r}_{\hat{C}_{7,j}}$	0	0	0	1,310,258	2,491,628	3,517,799	4,591,416	5,402,700	6,703,982	7,059,799
$\hat{C}_{8,j}$	376,686	1,363,294	2,382,128	3,471,744	4,075,313	4,498,426	4,886,502	5,149,760	5,544,000	5,642,266
$\hat{l}_{\hat{C}_{8,j}}$	0	0	1,018,834	2,108,450	2,712,019	3,135,132	3,523,208	3,786,466	4,180,706	4,278,972
$\hat{r}_{\hat{C}_{8,j}}$	0	0	1,018,834	3,040,506	4,701,269	6,100,587	7,541,250	8,617,067	10,330,671	10,795,153
$\hat{C}_{9,j}$	344,014	1,200,818	2,098,228	3,057,984	3,589,620	3,962,307	4,304,132	4,536,015	4,883,270	4,969,825
$\hat{l}_{\hat{C}_{9,j}}$	0	856,804	1,754,214	2,713,970	3,245,606	3,618,293	3,960,118	4,192,001	4,539,256	4,625,811
$\hat{r}_{\hat{C}_{9,j}}$	0	856,804	3,034,848	6,770,961	9,656,884	12,034,794	14,453,088	16,242,272	19,076,387	19,839,189

Table 5.3: Filled run-off-triangle with observed cumulative claims $C_{i,j}$ $(i+j \leq I)$ and predicted cumulative claims $\hat{\tilde{C}}_{i,j}$ $(i+j > I)$.

accident-year i	$\hat{\tilde{R}}_i = (\hat{R}_i, \hat{l}_{\hat{R}_i}, \hat{r}_{\hat{R}_i})$ $\hat{R}_i$	$\hat{l}_{\hat{R}_i}$	$\hat{r}_{\hat{R}_i}$
0	0.00	0.00	0.00
1	94,633.81	94,633.81	94,633.81
2	469,511.29	469,511.29	482,834.38
3	709,637.82	709,637.82	770,712.24
4	984,888.64	984,888.64	1,148,703.01
5	1,419,459.46	1,419,459.46	1,802,935.09
6	2,177,640.62	2,177,640.62	3,130,917.40
7	3,920,301.01	3,920,301.01	7,059,798.97
8	4,278,972.26	4,278,972.26	10,795,153.00
9	4,625,810.69	4,625,810.69	19,839,189.18
$\sum$	18,680,855.61	18,680,855.61	45,124,877.08

Table 5.4: Predicted reserves $\hat{\tilde{R}}_i$ for $i = 0, \ldots, I$.

zero. Since the spreads of the reserves are relatively wide the slope of the membership function for the left spread is very small. As a result, values close to zero are possible but only with a small grade of membership.

accident-year i	$E_\beta(\hat{\tilde{R}}_i)$ $\beta = 0.1$	$\beta = 0.25$	$\beta = 0.5$	$\beta = 0.75$	$\beta = 0.9$	chain-ladder reserve $\hat{R}_i$
0	0.00	0.00	0.00	0.00	0.00	0.00
1	56,780.29	70,975.36	94,633.81	118,292.27	132,487.34	94,633.81
2	282,372.93	353,798.85	472,842.06	591,885.27	663,311.20	469,511.29
3	428,836.41	539,862.67	724,906.43	909,950.18	1,020,976.44	709,637.82
4	599,123.90	759,143.28	1,025,842.23	1,292,541.19	1,452,560.56	984,888.64
5	870,849.46	1,112,529.05	1,515,328.37	1,918,127.68	2,159,807.27	1,419,459.46
6	1,354,248.21	1,752,390.06	2,415,959.81	3,079,529.57	3,477,671.42	2,177,640.62
7	2,509,155.51	3,332,663.00	4,705,175.50	6,077,688.00	6,901,195.50	3,920,301.01
8	2,893,192.39	4,023,751.79	5,908,017.45	7,792,283.11	8,922,842.50	4,278,972.26
9	3,536,155.34	5,371,030.33	8,429,155.31	11,487,280.30	13,322,155.29	4,625,810.69
$\sum$	12,530,714.44	17,316,144.39	25,291,860.98	33,267,577.57	38,053,007.52	18,680,855.61

Table 5.5: Expected values of the reserves $\hat{\tilde{R}}_i$, $i = 0, \ldots, I$, for different choices of the "decision-maker risk parameter" β and corresponding chain-ladder reserves $\hat{R}_i$, $i = 0, \ldots, I$.

Actuaries cannot set up a fuzzy magnitude as a reserve. Thus, expected values for the estimated reserves $E_\beta(\hat{\tilde{R}}_i)$, $i = 0, \ldots, I$, for different choices of the decision-maker risk parameter β, $\beta = 0.1, 0.25, 0.5, 0.75, 0.9$, are given in Table 5.5. The classical CL method and the fuzzy one proposed here cannot be compared, because the first considers uncertainty due to stochastic randomness whereas the latter takes into account vagueness due to subjective judgement. Nevertheless, the resulting reserves are shown in the very right column of Table 5.5. In fact, the "same value" of the estimated reserve for both the classical and the fuzzy method results for a choice of $\beta \approx 0.27$.[57] The higher the parameter β, the more weight is on the right spread resulting in a higher expected reserve for each specific accident year. Since the expected reserve for aggregated accident years is the sum of the expected reserves for single accident years the same observation holds true there.

The definition of Estimator 5.10 is chosen carefully due to the fact that the estimators are symmetrical and the left border is given by one. This leads to risk-averse and careful predictions of the estimated reserves. Beyond that, the use of secant approximation leads to more conservative predictions (cf. Remark 2.31c)).

[57]Note that a direct comparison is actually not applicable here.

accident-year i	$\mathrm{Unc}_K(\hat{\hat{C}}_{i,J} \mid \mathcal{D}_I)$ $K = 0.5$	$K = 1$	$K = 2$	$K = 5$	$K = 10$
0	0.00	0.00	0.00	0.00	0.00
1	47,316.91	94,633.81	189,267.63	473,169.07	946,338.15
2	238,086.42	476,172.84	952,345.67	2,380,864.18	4,761,728.35
3	370,087.51	740,175.03	1,480,350.06	3,700,875.15	7,401,750.29
4	533,397.91	1,066,795.82	2,133,591.65	5,333,979.12	10,667,958.25
5	805,598.64	1,611,197.27	3,222,394.55	8,055,986.36	16,111,972.73
6	1,327,139.50	2,654,279.01	5,308,558.02	13,271,395.05	26,542,790.09
7	2,745,025.00	5,490,049.99	10,980,099.98	27,450,249.96	54,900,499.92
8	3,768,531.32	7,537,062.63	15,074,125.26	37,685,313.16	75,370,626.32
9	6,116,249.97	12,232,499.94	24,464,999.87	61,162,499.68	122,324,999.35
$\sum$	15,951,433.17	31,902,866.35	63,805,732.69	159,514,331.73	319,028,663.46

Table 5.6: Prediction uncertainty $\mathrm{Unc}_K(\hat{\hat{C}}_{i,J} \mid \mathcal{D}_I)$ for each accident year $i = 0, \ldots, I$ for different choices of $K \in \mathbb{R}^+$.

The variety of the computed ultimate claims is given in Table 5.6 for several choices of the parameter K, $K = 0.5, 1, 2, 5, 10$. Due to the definition of the uncertainty (cf. Definition 2.50) the values are proportional to each other. For each accident year the ultimate claim is given by a triangular fuzzy number of which we find the area between the membership function and the x-axis for $K = 1$. The sum of the areas of the fuzzy numbers for each particular K is given in the bottom line. Intuitively, the uncertainty dependent on K is increasing for an increasing choice of the parameter K which is a weight factor of the computed areas as well as later accident years come along with higher uncertainty values since more fuzzy numbers need to be multiplied. Moreover, for rising i the predicted values are farer away from the present. This leads to an increasing uncertainty in the predictions.

We finish the chapter with a concluding discussion of the presented claims reserving method.

5.6 Discussion

The claims reserving method presented in the present chapter is an addition to the classical reserving techniques. While previous applications of fuzzy set theory to the field of loss reserving only made use of fuzzy regression (cf. Section 4.6). Hence, the FCL method can be seen as a first attempt to apply fuzzy numbers and its corresponding arithmetic to the field of claims reserving. It goes back to the roots of the CL method which initially has been a purely deterministic reserving principle.

As elaborated on in Section 5.1 chain-ladder factors are sometimes modified retrospectively in practice due to the personal experience of the reserving actuary. This action adds vagueness and imprecision to the development factors which is here captured by the triangular fuzzy numbers. Thus, no later alteration is needed. Furthermore, with the uncertainty measure Unc_K a means to quantify the vagueness of the calculated reserves is given.

The choice of the FCL factors as triangular fuzzy numbers is rather simple but leads to easy to conduct computations. The selection results in predicted reserves with a relatively wide spread. Needless to say, the presented method is not limited to the use of triangular fuzzy numbers. Membership functions with e.g. an exponential shape (cf. Definition 2.21) are also thinkable here. They would assign a higher grade of membership to the values close to the classical CL estimate.
Additionally, Estimator 5.5 leads to a left spread ranging to one, but this is not mandatory. Another possible extension might be the use of asymmetric FN in order to cope with stronger deviations in only one direction. One possible application is presented in Chapter 6 in combination with fuzzy regression.

On the other hand, the FCL model only considers fuzziness whereas stochastic randomness is not regarded. Hence, classical and fuzzy methods should not be compared without taking into account that it is dealt with different sources of uncertainty (cf. Section 2.1.4). A possible solution could be given by an application of fuzzy random variables which can handle both sources of uncertainty (cf. Section 2.4).

The values of the uncertainty measure for increasing accident years as well as for a higher choice of the parameter K seem to go up rapidly. At this point of research there is no proper interpretation of these quantities on their own. It is only possible to compare uncertainties of the ultimate claims for different run-off triangles among themselves. By the use of fuzzy numbers there is no balancing of the uncertainty of the ultimate claims over several accident years.

6 Another Fuzzy Chain-Ladder Model – An Application of Fuzzy Regression Techniques

In this chapter an application of fuzzy regression methods to the chain-ladder method is presented. Firstly, a motivation as well as a description of the notation and setting is given in Section 6.1. In Section 6.2 we propose another fuzzy chain-ladder model making use of fuzzy regression as presented by Ishibuchi and Nii (2001). Subsequently, an example is conducted in Section 6.3 where the results are shown in Section 6.3.1 and are then compared to those of the FCL method (cf. Chapter 5) in Section 6.3.2. Finally, the proposed model is discussed in Section 6.4.

6.1 Introduction

6.1.1 Motivation

As elaborated on in Chapter 1 there are many facts which justify the utilization of fuzzy set theory in claims reserving. As one of the main reasons subjective judgment can be argued.

Depending on the line of business (LoB) there may be only small datasets available in order to determine the claims reserve. Moreover, while determining the development factors with the classical chain-ladder method the cardinality of the set of observations which underlie the calculation decreases with increasing development years. In the same way as in the FCL method (cf. Chapter 5) the vagueness is expressed by triangular fuzzy numbers which are advantageous since they are easy to handle arithmetically and convey major parts of the uncertainty. The use of symmetrical fuzzy numbers can come along with some disadvantages as seen in the FCL method. In the latter procedure the estimation of the CL-factors is chosen in a manner that the left spread adjoins one (cf. Remarks 5.6). As the left and right spreads are chosen to be of equal length, this might result in a large total spread. As a consequence, the result is rather conservative. In order to circumnavigate this drawback we here utilize asymmetrical fuzzy numbers which allow for an individual adaptation of the left and right spread depending on the situation.

The classical chain-ladder method can be perceived as a sequence of weighted regressions from one development year to the next one (cf. p. 78). Then, the CL-factors are yielded with the help of the Aitken-estimator (cf. Equation (A.1)). Hence, the sequence of weighted linear regressions can also be transferred to a sequence of ordinary linear regressions by multiplying the dependent and independent variables with the corresponding weighing matrix. Thus, the CL-factors are simply the Gauß-Markov estimators in this case.

Fuzzy regression techniques often make use of a H-value (or h-certain factor) which can be interpreted as the confidence the user has in the data (cf. Section 2.3). An extensive discussion of the usage of this value can be found in Moskowitz and Kim (1993). Thus, we propose a model in which the h-certain factor can be adjusted for every single regression in the sequence of regressions. By providing a pattern which is a vector of h-certain factors the trust in the data for each single regression is specified. As the data situation is better for earlier development years we propose smaller values of the h-certain factor for earlier development years. The data situation is understood to be better in the sense that more observations are available for earlier development years. Often one can argue that the values in the pattern will be monotonically increasing.

So far, fuzzy regression techniques have seldom been applied to the chain-ladder reserving method in the literature though CL is still the most popular method. As stated in Section 4.6 fuzzy regression models have been used as a modification of

Sherman's scheme, the method of Benjamin and Eagles, Taylor's separation method and Kremer's two ways of ANOVA.

One of the rare publications considering the CL-method in a fuzzy setup by Kerkez propose a reserving model in which a fuzzy regression method by Tanaka et al. (1982) is applied to the chain-ladder method (cf. Section 4.6.6). Kerkez sketches that some reserving methods (among others CL) can be perceived as linear regression models which is the motivation for her model. The author makes use of symmetrical fuzzy numbers as coefficients such that a large deviation in only one direction results in the same width of the spread in the other direction. An example of the reserving method is conducted both with and without intercept where the latter case serves as a comparison for the chain-ladder method. Kerkez does not make use of a defuzzification method in order to determine a crisp reserve. In fact, the model yields fuzzy reserves.[58] The publication by Andrés Sánchez and Terceño Gómez (2003) obtains the CL method as a special case.

Some of the earlier published fuzzy regression models have some serious drawbacks (cf. Section 2.3). E.g. the fuzzy regression coefficients are chosen to be symmetrical and, thus, a strong deviation in one direction results in a large total spread of the coefficients. In order to avoid these disadvantages Ishibuchi and Nii proposed a model with asymmetric coefficients (cf. Section 2.3.1 and Ishibuchi and Nii 2001).

Even though the CL method is the most popular reserving method it has not been in the core of research for fuzzy applications. The goal of this chapter is to define a fuzzy CL-model which counterbalances some of the drawbacks. Therefore, the fuzzy regression model of Ishibuchi and Nii (2001) with asymmetric coefficients is applied. Then, the total spread of the fuzzy coefficients is not needlessly large. Compared to the FCL model as introduced in Chapter 5, the model goes without strong restrictions on the spreads of the development factors whose supports are bounded below by one. This results in smaller total spreads here since deviations are captured in a better way and, thus, a smaller uncertainty of the predicted reserve.

[58]Onoghojobi and Olewuezi (2015) also apply fuzzy regression techniques to the CL method. As elaborated on before (cf. p. 105) we will not present the findings of the publication in detail.

6.1.2 Notation and Methods

The model is based on a development triangle as shown in Figure 4.2 which contains cumulative claims $C_{i,j}$, $i \in \{0, \dots, I\}$, $j \in \{0, \dots, J\}$, made in relative accident year i and relative development year j. We assume that it is a real development triangle, i.e. $I = J$, and that the cumulative claims are strictly positive.

The method presented here makes use of TFNs as introduced in Definition 2.18, their corresponding arithmetic and the fuzzy regression model by Ishibuchi and Nii (2001) (see Section 2.3.1). The secant approximation (cf. Equation (2.7)) is used for the multiplication of TFNs. As defuzzification method the expected value as defined in Definition 2.38 is used. In order to quantify the prediction uncertainty the measure of uncertainty as defined in Definition 2.50 is applied.

6.2 The Model

In Chapter 5 it has been pointed out that there might be situations in which actuaries adjust the development factors calculated by the CL-method according to their subjective judgment. This observation is also taken into account in this chapter. Precisely speaking, the proposed model here is not a means to picture stochastic randomness but deals with vagueness.

The goal is to predict claims reserves. Hereby, the claims development is considered with the help of development factors. Those shall be estimated here in the context of a fuzzy CL-method based on a fuzzy regression model. Therefore, the fuzzy CL-factors $\tilde{f}_j^{\mathrm{AFCL}} = \left(f_j^{\mathrm{AFCL}}, l_{f_j^{\mathrm{AFCL}}}, r_{f_j^{\mathrm{AFCL}}}\right)$, $j \in \{0, \dots, J-1\}$, are estimated with fuzzy regression methods. Hereby, the modes of the estimated development factors $\hat{\tilde{f}}_j^{\mathrm{AFCL}} = \left(\hat{f}_j^{\mathrm{AFCL}}, \hat{l}_{\hat{f}_j^{\mathrm{AFCL}}}, \hat{r}_{\hat{f}_j^{\mathrm{AFCL}}}\right)$, $j \in \{0, \dots, J-1\}$, are estimated with OLS. The CL-method can be stated in the context of linear models. In fact, every single development factor is derived as the estimated regression coefficient in a general multiple linear regression model without intercept (cf. p. 78). By choosing appropriate vectors $\mathbf{y}^*$, $\mathbf{X}^*$ and $\boldsymbol{\varepsilon}^*$ the weighted multiple linear model can be transformed in a classical one with the help of a weighting matrix $\mathbf{W}$ (cf. Equations (4.9)-(4.11)). In order to apply the regression method by Ishibuchi and Nii (2001) we are here regarding the transformed data $\mathbf{y}^*$, $\mathbf{X}^*$

and $\boldsymbol{\varepsilon}^*$ as defined in Equations (4.9)-(4.10) to derive estimators for the fuzzy CL-factors $\tilde{f}_j^{\text{AFCL}}$, $j \in \{0, \dots, J-1\}$.

According to the definitions in Ishibuchi and Nii (2001) (see also Section 2.3.1) we propose the following fuzzy regression model without intercept.

Model Assumptions 6.1 (Fuzzy Chain-Ladder Regression Model)
Let $\tilde{f}_j^{\text{AFCL}}$, $j \in \{0, \dots, J-1\}$, be a triangular fuzzy number given by $\tilde{f}_j^{\text{AFCL}} = \big(f_j^{\text{AFCL}}, l_{f_j^{\text{AFCL}}}, r_{f_j^{\text{AFCL}}}\big)$, $\widetilde{\mathbf{Y}}^* = (\tilde{y}_0^*, \dots, \tilde{y}_{J-j-1}^*)^T$ a fuzzy output and $\mathbf{X}^* = (x_0^*, \dots, x_{J-j-1}^*)^T$ a non-fuzzy input vector with $x_0^*, \dots, x_{J-j-1}^*$ being strictly positive. Then, the fuzzy chain-ladder regression model for the j-th development factor is given by

$$\widetilde{\mathbf{Y}}^* = \tilde{f}_j^{\text{AFCL}} \otimes \mathbf{X}^*.$$

We refer to $\tilde{f}_j^{\text{AFCL}}$ as the fuzzy coefficient and the multiplication needs to be understood component wise.

Remarks 6.2

a) We assume that cumulative figures are being used in the development triangle which are strictly positive. Then, the weighting matrix $\mathbf{W}$ (cf. Equation (4.7)) is positive definite and, thus, invertible. That way it is assured that $\mathbf{W}^{-1/2}$ exists and thereby the transformed variables can be calculated.

b) When considering the above fuzzy regression model line by line we yield for a single entry in the estimated fuzzy output vector:

$$\begin{aligned}
\hat{\tilde{y}}_i^* &= \hat{\tilde{f}}_j^{\text{AFCL}} \otimes x_i \\
&= \left(\hat{f}_j^{\text{AFCL}}, \hat{l}_{f_j^{\text{AFCL}}}, \hat{r}_{f_j^{\text{AFCL}}}\right) \otimes x_i \\
&= \left(\hat{f}_j^{\text{AFCL}} x_i, \begin{cases} \hat{l}_{f_j^{\text{AFCL}}}|x_i|, & \text{if } x_i \geq 0 \\ \hat{r}_{f_j^{\text{AFCL}}}|x_i|, & \text{if } x_i < 0 \end{cases}, \begin{cases} \hat{r}_{f_j^{\text{AFCL}}}|x_i|, & \text{if } x_i \geq 0 \\ \hat{l}_{f_j^{\text{AFCL}}}|x_i|, & \text{if } x_i < 0 \end{cases}\right) \\
&= \left(\hat{f}_j^{\text{AFCL}} x_i, \hat{l}_{f_j^{\text{AFCL}}} x_i, \hat{r}_{f_j^{\text{AFCL}}} x_i\right), \qquad i = 0, \dots, I-j-1
\end{aligned}$$

The simplification in the last above equation is due to the fact that strictly positive figures are assumed.

The determinations of the mode as well as the left and right spread of the fuzzy coefficient go in line with the proposition of Ishibuchi and Nii (2001) (cf. the estimation method described on p. 54 and especially the linear program in Equation (2.14)). The mode f_j^{AFCL} is specified by using OLS and we refer to the estimation of f_j^{AFCL} as $\hat{f}_j^{\text{AFCL}}$. The estimated left and right spreads $\hat{l}_{f_j^{\text{AFCL}}}$ and $\hat{r}_{f_j^{\text{AFCL}}}$, respectively, are determined by an optimization problem in which the total fuzziness of the model is minimized. Let $\boldsymbol{\alpha}^* = (\alpha_0^*, \ldots, \alpha_{J-1}^*)$ be a vector where $\alpha_j^* \in (0,1)$ for all $j \in \{0, \ldots, J-1\}$. Then, the linear program for the determination of the spreads of a single estimated development factor $\hat{\tilde{f}}_j^{\text{AFCL}}$ for a specific j is given by:

$$\min_{\hat{l}_{f_j^{\text{AFCL}}}, \hat{r}_{f_j^{\text{AFCL}}}} \sum_{i=0}^{I-j-1} \hat{l}_{f_j^{\text{AFCL}}} x_i + \sum_{i=0}^{I-j-1} \hat{r}_{f_j^{\text{AFCL}}} x_i \tag{6.1a}$$

s.t.

$$\hat{f}_j^{\text{AFCL}} x_i - \hat{l}_{f_j^{\text{AFCL}}} x_i (1-\alpha_j^*) \leq y_i, \qquad i = 0, \ldots, I-j-1 \tag{6.1b}$$

$$\hat{f}_j^{\text{AFCL}} x_i + \hat{r}_{f_j^{\text{AFCL}}} x_i (1-\alpha_j^*) \geq y_i, \qquad i = 0, \ldots, I-j-1 \tag{6.1c}$$

$$\hat{l}_{f_j^{\text{AFCL}}}, \hat{r}_{f_j^{\text{AFCL}}} \geq 0 \tag{6.1d}$$

Constraints (6.1b) and (6.1c) ensure that the – in this case – crisp output lies within the support of the estimated fuzzy output vector whereas constraint (6.1d) ensures that the spreads of the fuzzy coefficient are non-negative.

Literature on fuzzy regression suggests that the above linear program (6.1) can be solved with methods of linear optimization like e.g. the simplex algorithm. However, the linear program can be simplified in the following way when keeping in mind that only two of the equations listed in Equations (6.1b) and (6.1c) are binding. Equations (6.1b) and (6.1c) can be rewritten as follows:

$$\begin{aligned}
(6.1\text{b}) \Leftrightarrow \quad & \hat{l}_{f_j^{\text{AFCL}}} \geq \frac{\hat{f}_j^{\text{AFCL}}}{(1-\alpha_j^*)} - \frac{y_i}{x_i(1-\alpha_j^*)}, && i = 0, \ldots, I-j-1 \\
\Leftrightarrow \quad & \hat{l}_{f_j^{\text{AFCL}}} \geq \max_{i=0,\ldots,I-j-1} \left(\frac{\hat{f}_j^{\text{AFCL}}}{(1-\alpha_j^*)} - \frac{y_i}{x_i(1-\alpha_j^*)} \right) && (6.2) \\
& \qquad = \max_{i=0,\ldots,I-j-1} \frac{1}{(1-\alpha_j^*)} \left(\hat{f}_j^{\text{AFCL}} - \frac{y_i}{x_i} \right) \\
(6.1\text{c}) \Leftrightarrow \quad & \hat{r}_{f_j^{\text{AFCL}}} \geq \frac{y_i}{x_i(1-\alpha_j^*)} - \frac{\hat{f}_j^{\text{AFCL}}}{(1-\alpha_j^*)}, && i = 0, \ldots, I-j-1
\end{aligned}$$

$$\Leftrightarrow \quad \hat{r}_{f_j^{\text{AFCL}}} \geq \max_{i=0,\ldots,I-j-1} \left(\frac{y_i}{x_i(1-\alpha_j^*)} - \frac{\hat{f}_j^{\text{AFCL}}}{(1-\alpha_j^*)} \right) \tag{6.3}$$
$$= \max_{i=0,\ldots,I-j-1} \frac{1}{(1-\alpha_j^*)} \left(\frac{y_i}{x_i} - \hat{f}_j^{\text{AFCL}} \right)$$

Consequently, the restrictions are only binding for the largest and smallest individual development factor $\frac{y_i}{x_i}$. Solely the difference between the largest (respectively smallest) individual development factor and the estimated CL-factor is considered.
By choosing the elements of the vector $\boldsymbol{\alpha}^*$ which are the H-values for every single fuzzy regression model the actuary expresses his/her confidence in the data. In the special case of selecting a h-certain factor α_j^* as zero the definition of the applied model leads to the interpretation that those data points which mark the left and right spread, i.e. which refer to the binding equations in the minimization problem (6.1a)-(6.1d), are of less importance since they are assessed with a grade of membership of zero. If an actuary is of the opinion that these "more extreme" cases should be considered as well he/she can attach more importance to them by choosing a higher value for α_j^*. Therefore, the elements of the vector $\boldsymbol{\alpha}^*$ can be as well interpreted as an assessment of the actuary if the usual scope of values regarding their variability (with respect to the considered LoB) is covered by the regression "tube" or not.

While there normally is sufficient data in earlier development years, the data is getting sparser for increasing development years due to the triangular structure. This results in the extreme that there is only one single data point available when calculating the FCL-factor $\hat{\hat{f}}_{J-1}$. Whereas the mode can still be determined with OLS there are no reasonable results for the spreads. Generally, the width of the left and right spread possess a decreasing behavior. This is due to the fact that for higher development years more claims are finally settled. Accordingly, the CL-factor approaches one and the variability generally decreases. In this way they resemble the characteristics of the variance parameters in the classical CL-method. Hence, we propose to extrapolate likewise the left and right spread of the last FCL-factor by log-linear regression as mentioned in Mack (1993, p. 217).[59]
We consider the sequence of the logarithms of the spreads $\ln \hat{l}_{f_0^{\text{AFCL}}}, \ldots, \ln \hat{l}_{f_{J-2}^{\text{AFCL}}}$ as dependent and the sequence of development years $0, \ldots, J-2$ as independent variables,

[59] There exists a broad literature on the estimation/extrapolation of variance parameters. Other approaches like e.g. by assuming that the ratios of the two previous variance parameters are alike, i.e. $\frac{\hat{\sigma}_{J-3}}{\hat{\sigma}_{J-2}} = \frac{\hat{\sigma}_{J-2}}{\hat{\sigma}_{J-1}}$, can be equally transferred.

i.e. in the notation of OLS $\mathbf{y} := (\ln \hat{l}_{f_0^{\text{AFCL}}}, \ldots, \ln \hat{l}_{f_{J-2}^{\text{AFCL}}})^T$ and $\mathbf{x} := (0, \ldots, J-2)^T$. We fit a line to the given data points, i.e.

$$\mathbf{y} = a_0 + a_1 \mathbf{x}$$

with unknown coefficients $a_0, a_1 \in \mathbb{R}$ whose estimates $\hat{a}_0$ and $\hat{a}_1$ are derived with OLS (cf. Appendix A). Thus, an estimate of the left spread of the fuzzy chain-ladder factor for the development year $J-1$ is given by $\hat{l}_{f_{J-1}} = \exp(\hat{a}_0 + \hat{a}_1(J-1))$ by extrapolation. With the same procedure an estimation of the right spread of the $(J-1)$-th development factor is derived.

As previously elaborated on in Chapter 5 the determination of the prediction for the ultimate claims is carried out by gradually fuzzy multiplying the last observation for a given accident year with the estimated fuzzy development factors. Beyond that, predictions for the reserves for both single and aggregated accident years can be deduced with the help of fuzzy arithmetic as well.

6.3 Example

For the sake of comparison we consider the same data set as in Section 5.5 given in Table 5.1. We also apply the same measure of uncertainty as in Chapter 5. In this section we will firstly present the results derived with the proposed method in this chapter.[60] Subsequently, the findings will be compared to those of the FCL method (cf. Section 5.5) in Section 6.3.2.

6.3.1 Presentation of Results

We consider the development triangle as given in Table 5.1. For the parameter vector $\boldsymbol{\alpha}^* = (\alpha_0, \ldots, \alpha_8)$ we assume the following pattern as shown in Table 6.1. The value of the h-certain factor increases for later development years. Therefore, in this case,

[60] The left and right spreads are calculated directly using Equations (6.2) and (6.3). The same results (for this particular example) are derived when the optimization problems are solved using the library “linprog” of the R-package “LPsolve” for linear problems in R. R is an open source statistics software.

there is higher confidence that the data for earlier development years already show a sufficient picture of the development factors to estimate whereas the information of the development years six and higher is judged as not sufficient. Hence, the actuary chooses a "safety margin" in form of parameters α_j others than zero for development years six and higher, i.e. $\alpha_j > 0$, $j = 5, \ldots, 8$.

	development year j								
	0	1	2	3	4	5	6	7	8
α_j	0	0	0	0	0	0.2	0.4	0.6	0.8

Table 6.1: An α-pattern $\boldsymbol{\alpha}^* = (\alpha_0, \ldots, \alpha_8)$.

The modes of the estimated fuzzy chain-ladder factors are the estimated CL-factors (cf. Table 5.2).

A pattern of this particular type is picked in situations in which the actuary has high confidence in the data. From the optimization problem's point of view it means that the minimal spread such that all data points are included is calculated for the first five development years. For development years six and higher an addition to the spread is warranted. This could be due to the fact that there are less data points available and in the actuary's eyes these might not reflect a comprehensive picture of possible spreads. This impression could be even reinforced for higher development years such that increasing values for α_j are chosen in this particular case. In this situation, we yield the following results as displayed in Table 6.2.

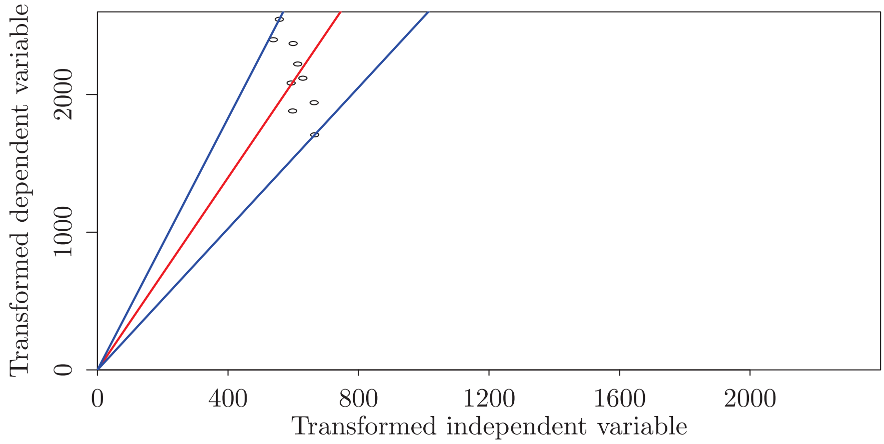
accident-year i	development year j								
	0	1	2	3	4	5	6	7	8
$\hat{\tilde{f}}_1^{\text{AFCL}}$	3.491	1.747	1.457	1.174	1.104	1.086	1.054	1.077	1.018
$\hat{l}_{f_j^{\text{AFCL}}}$	0.926	0.205	0.179	0.101	0.067	0.053	0.023	0.034	0.013
$\hat{r}_{f_j^{\text{AFCL}}}$	1.077	0.268	0.254	0.064	0.105	0.052	0.011	0.025	0.008

Table 6.2: Calculated FCL-factors for $\boldsymbol{\alpha}^* = (0, 0, 0, 0, 0, 0.2, 0.4, 0.6, 0.8)$.

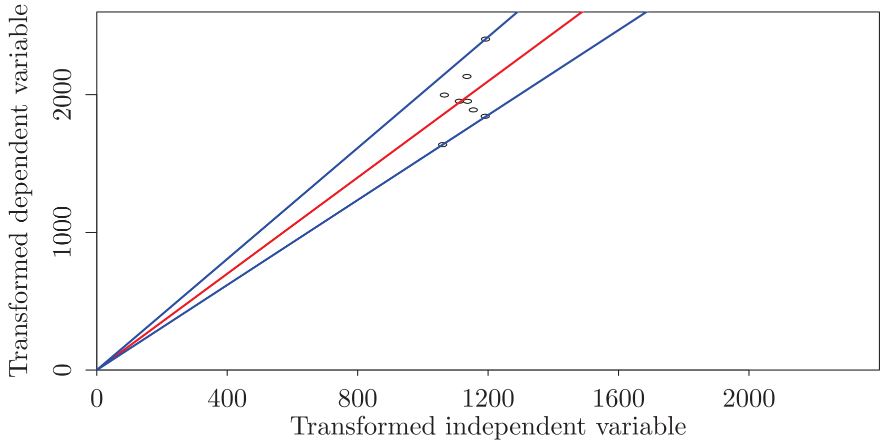
Again, we write down the fuzzy chain-ladder factors in three lines. The first line indicates the mode whereas in the second and third line the left and right spread, respectively, is given. Due to the estimation procedure the modes coincide with the classical chain-ladder factors. In contrast to the values given in Table 5.2 the estimates of the left and right spread in Table 6.2 are not equal since asymmetric coefficients are assumed. As anticipated, the left and right spreads show a decreasing behavior with only slight variations. While for the development factor zero the left and right spreads are given by 0.926 and 1.077 the width of the spreads go down for increasing development years with only one exception (development year seven) and are given by 0.034 and 0.025 for the seventh development factor. Therefore, this example confirms the anticipated behavior. The exception in development year 7 is due to the fact that the individual CL-factors are higher for that specific development year. The spreads of the eighth development factor are yielded by extrapolation by means of log-linear regression (see Table 6.2). The plotted spreads as well as the fitted exponential functions to the data points are shown in Figure 6.4. The case of the left spreads is visualized in Figure 6.4a whereas Figure 6.4b depicts the right spreads. Both figures show that the estimated exponential function fits well the data points for higher development years. This is supported by the fact that the conducted regression gives a coefficient of determination of $R^2 \approx 0.8924$ and an adjusted value of $R_a^2 \approx 0.8745$. For the left and right spreads there are a little greater deviations for earlier development years between the data points and the fitted exponential function, but the discrepancies diminish for later development years. For the right spreads the regression still yields $R^2 \approx 0.8677$ and $R_a^2 \approx 0.8456$ and therefore it militates in favor of a good fit.

Figures 6.1, 6.2 and 6.3 show the regression "tubes" which result with the fuzzy regression approach. For illustration the transferred cumulated claims are plotted. Then, the red line is the regression line yielded by OLS of which the slope parameter is the chain-ladder factor. The blue lines are derived by the optimization problem for the vector $\boldsymbol{\alpha}^*$ as given in Table 6.1 where the slope parameter are the left and right spread, respectively.

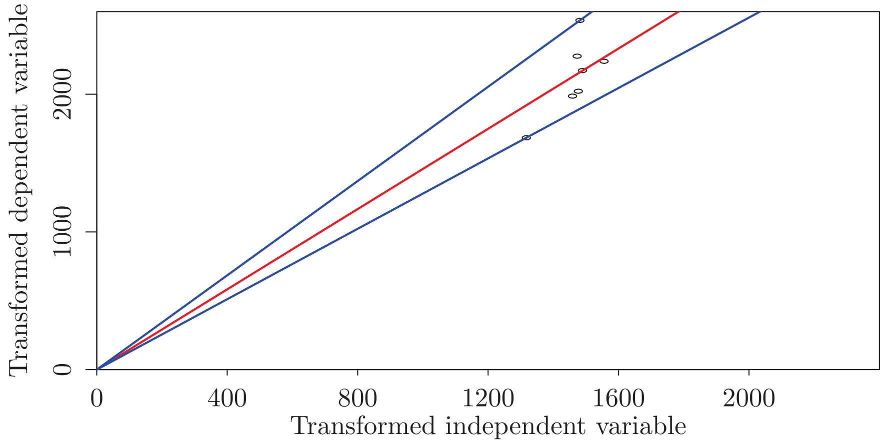
The fuzzy regression of the development year 0 to 1 , i.e. the column 0 to the first column of the run-off triangle, is visualized in Figure 6.1a. As there are greater upward deviations, the regression "tube" has a larger right spread (shown by the upper blue straight line). This fact is captured by the asymmetric coefficient. For the earlier development years the h-certain factors are chosen as zero. In this case, the blue lines (for development years one to five) always pass through those data points which have the maximal distance to the regression line derived by OLS. For later development years another factor is chosen as e.g. $\alpha_6 = 0.4$. This results in larger spreads of the



(a) Fuzzy regression of zero to first development year.

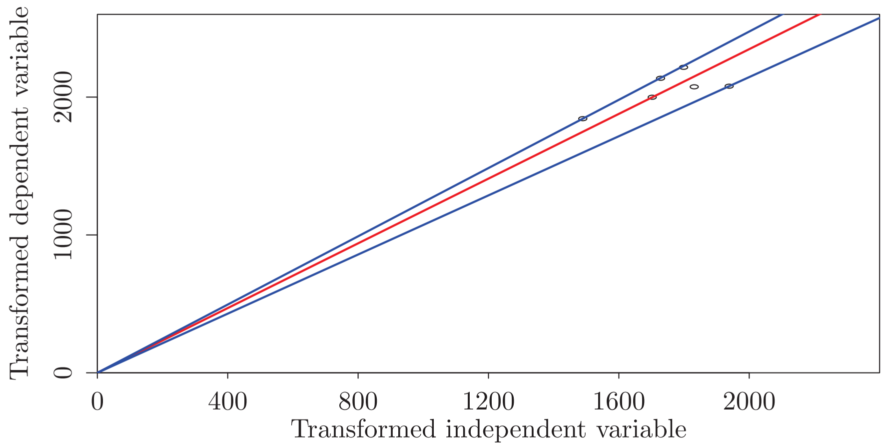


(b) Fuzzy regression of first to second development year.

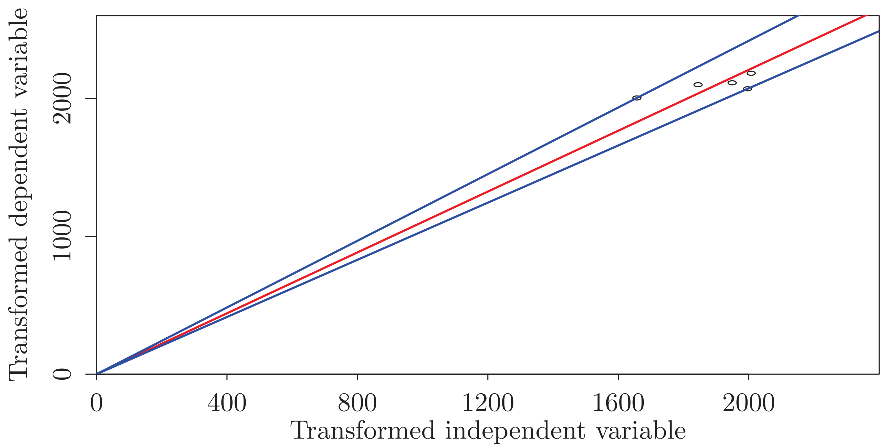


(c) Fuzzy regression of second to third development year.

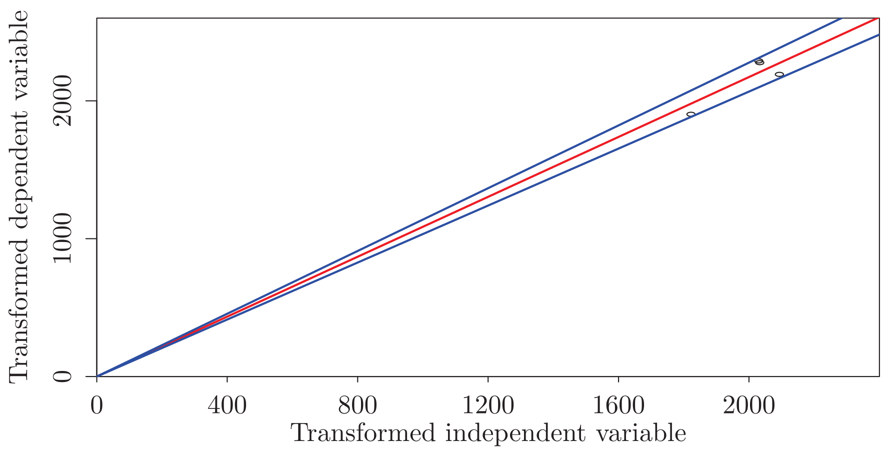
Figure 6.1: Illustration of regression lines for fuzzy chain-ladder factors $\hat{\hat{f}}_0$, $\hat{\hat{f}}_1$ and $\hat{\hat{f}}_2$.



(a) Fuzzy regression of third to fourth development year.

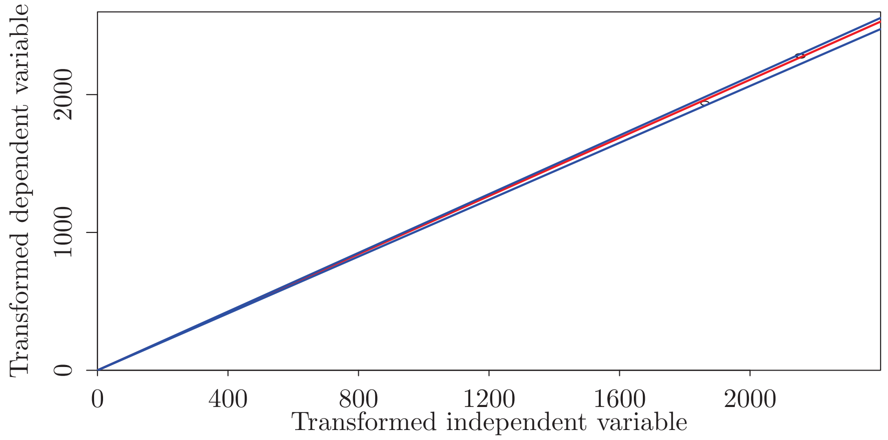


(b) Fuzzy regression of fourth to fifth development year.

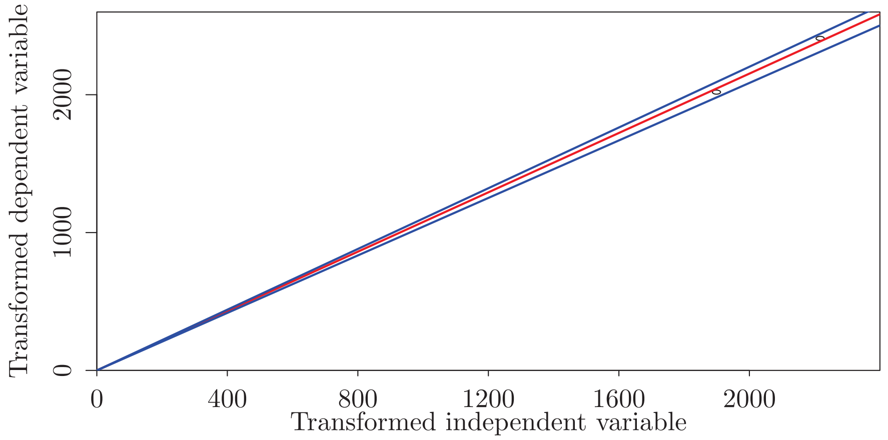


(c) Fuzzy regression of fifth to sixth development year.

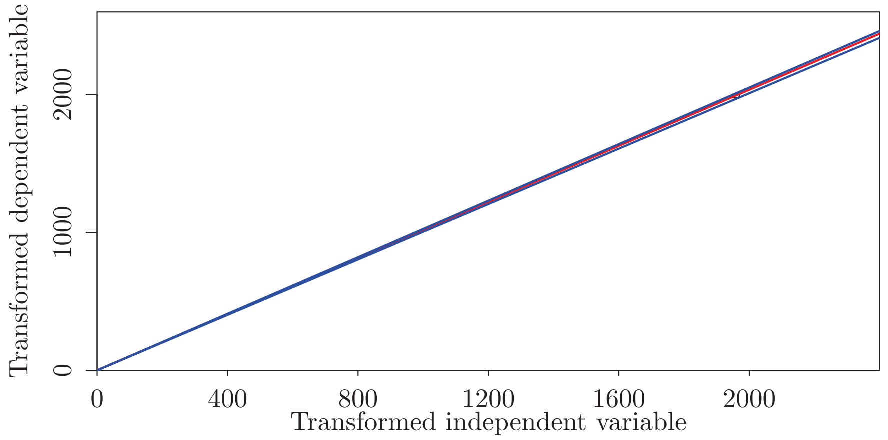
Figure 6.2: Illustration of regression lines for fuzzy chain-ladder factors $\hat{\hat{f}}_3$, $\hat{\hat{f}}_4$ and $\hat{\hat{f}}_5$.



(a) Fuzzy regression of sixth to seventh development year.



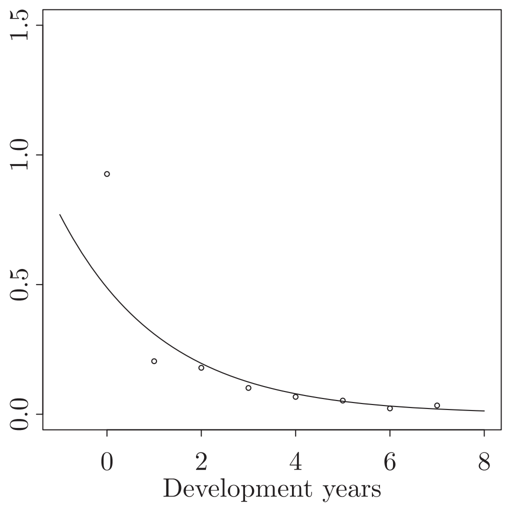
(b) Fuzzy regression of seventh to eighth development year.



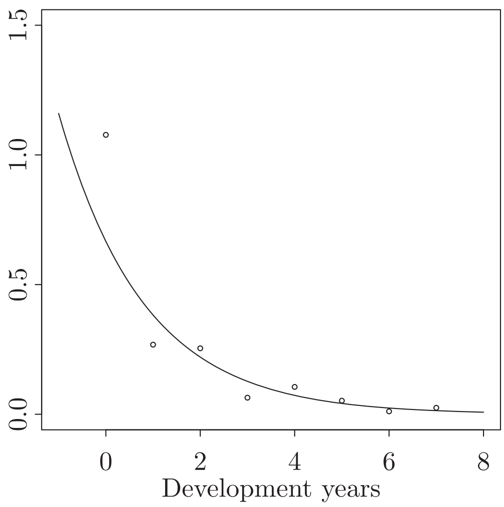
(c) Fuzzy regression of eighth to ninth development year.

Figure 6.3: Illustration of regression lines for fuzzy chain-ladder factors $\hat{\tilde{f}}_6$, $\hat{\tilde{f}}_7$ and $\hat{\tilde{f}}_8$.

fuzzy coefficients which allow for a higher variability (cf. Figure 6.3b). As seen in Table 6.2 the spreads decrease for increasing development years. Figure 6.3c shows the regression “tube” for which the left and right spreads are yielded by extrapolation by means of log-linear regression as there are too few data points.



(a) Extrapolation of left spread.



(b) Extrapolation of right spread.

Figure 6.4: Left and right spreads as well as the fitted exponential functions by log-linear regression.

The filled development triangle with the help of the fuzzy chain-ladder factors as given in Table 6.2 is shown in Table 6.3. As in the previous Chapter 5 all cumulative claims as well as the predicted cumulative claims are written down in three lines for each specific accident and development year. The observed magnitudes are displayed with a gray background whereas the predicted magnitudes are shown with a white one.

The estimated reserves $\hat{\tilde{R}}_i$ for accident years i, $i \in \{0, \ldots, I\}$, are given in Table 6.4. For increasing accident years the modes as well as the left and right spreads of the predicted reserves are apparently rising. The aggregated reserve adds up to approximately 18.7 million with a left and right spread of approximately 10.7 million and approximately 16.8 million, respectively.

The reserves need to be specified as a crisp number. Therefore, Table 6.5 gives an overview of the expected values for several choices of the “decision-maker risk parameter” β, $\beta \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$. If one keeps in mind that the classical method and the

	development year j									
$\hat{\hat{C}}_{i,j}$	0	1	2	3	4	5	6	7	8	9
$\hat{C}_{0,j}$	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
$\hat{l}_{\hat{C}_{0,j}}$	0	0	0	0	0	0	0	0	0	0
$\hat{r}_{\hat{C}_{0,j}}$	0	0	0	0	0	0	0	0	0	0
$\hat{C}_{1,j}$	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	5,433,719
$\hat{l}_{\hat{C}_{1,j}}$	0	0	0	0	0	0	0	0	0	68,019
$\hat{r}_{\hat{C}_{1,j}}$	0	0	0	0	0	0	0	0	0	42,374
$\hat{C}_{2,j}$	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315	5,285,148	5,378,826
$\hat{l}_{\hat{C}_{2,j}}$	0	0	0	0	0	0	0	0	166,254	234,415
$\hat{r}_{\hat{C}_{2,j}}$	0	0	0	0	0	0	0	0	122,010	167,086
$\hat{C}_{3,j}$	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268	4,835,458	5,205,637	5,297,906
$\hat{l}_{\hat{C}_{3,j}}$	0	0	0	0	0	0	0	103,238	271,399	339,071
$\hat{r}_{\hat{C}_{3,j}}$	0	0	0	0	0	0	0	51,255	176,627	222,474
$\hat{C}_{4,j}$	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311	4,207,459	4,434,133	4,773,589	4,858,200
$\hat{l}_{\hat{C}_{4,j}}$	0	0	0	0	0	0	204,271	305,349	468,547	531,697
$\hat{r}_{\hat{C}_{4,j}}$	0	0	0	0	0	0	202,557	262,733	399,577	447,716
$\hat{C}_{5,j}$	396,132	1,333,217	2,180,715	2,985,752	3,691,712	4,074,999	4,426,546	4,665,023	5,022,155	5,111,171
$\hat{l}_{\hat{C}_{5,j}}$	0	0	0	0	0	247,211	470,407	584,765	767,710	835,519
$\hat{r}_{\hat{C}_{5,j}}$	0	0	0	0	0	389,046	656,058	748,180	939,990	1,003,969
$\hat{C}_{6,j}$	440,832	1,288,463	2,419,861	3,483,130	4,088,678	4,513,179	4,902,528	5,166,649	5,562,182	5,660,771
$\hat{l}_{\hat{C}_{6,j}}$	0	0	0	0	352,959	639,762	899,230	1,037,752	1,257,023	1,334,150
$\hat{r}_{\hat{C}_{6,j}}$	0	0	0	0	222,460	699,879	1,032,877	1,154,826	1,400,340	1,480,419
$\hat{C}_{7,j}$	359,480	1,421,128	2,864,498	4,174,756	4,900,545	5,409,337	5,875,997	6,192,562	6,666,635	6,784,799
$\hat{l}_{\hat{C}_{7,j}}$	0	0	0	513,072	973,323	1,337,358	1,667,479	1,852,008	2,140,782	2,236,385
$\hat{r}_{\hat{C}_{7,j}}$	0	0	0	728,647	1,168,492	1,929,386	2,479,615	2,706,543	3,134,909	3,268,265
$\hat{C}_{8,j}$	376,686	1,363,294	2,382,128	3,471,744	4,075,313	4,498,426	4,886,502	5,149,760	5,544,000	5,642,266
$\hat{l}_{\hat{C}_{8,j}}$	0	0	278,830	783,100	1,191,694	1,508,518	1,796,339	1,962,646	2,220,828	2,302,528
$\hat{r}_{\hat{C}_{8,j}}$	0	0	365,797	1,232,111	1,746,740	2,541,641	3,129,071	3,387,189	3,858,662	4,001,680
$\hat{C}_{9,j}$	344,014	1,200,818	2,098,228	3,057,984	3,589,620	3,962,307	4,304,132	4,536,015	4,883,270	4,969,825
$\hat{l}_{\hat{C}_{9,j}}$	0	318,697	737,287	1,318,296	1,723,773	2,027,685	2,304,640	2,473,791	2,733,010	2,808,846
$\hat{r}_{\hat{C}_{9,j}}$	0	370,639	1,069,280	2,364,107	3,121,408	4,152,715	4,935,346	5,304,448	5,955,093	6,146,665

Table 6.3: Filled run-off-triangle for $\boldsymbol{\alpha} = (0, 0, 0, 0, 0, 0.2, 0.4, 0.6, 0.8)$.

fuzzy one try to model different uncertainties but still likes to compare the aggregated reserves, they equal each other for $\beta \approx 0.389$.

For a choice of the parameter K of $K = 1$ the uncertainty $\mathrm{Unc}_K(\hat{\hat{C}}_{i,J} \mid \mathcal{D}_I)$ of the ultimate claims for all accident years given the observations $\mathcal{D}_I$ is calculated and displayed in Table 6.6. The aggregated uncertainty amounts up to approximately 13.7 million. Uncertainties for other choices of the parameter K can be derived by multiplying by the constant factor.[61]

[61] As stated before, the constant K is not obligatory in the definition of fuzzy uncertainty (cf. p. 45). Without any restrictions the constant K can be set to one. With the help of this constant different approaches of decision-makers can be incorporated.

accident-year i	$\hat{\tilde{R}}_i = (\hat{R}_i, \hat{l}_{\hat{R}_i}, \hat{r}_{\hat{R}_i})$ $\hat{R}_i$	$\hat{l}_{\hat{R}_i}$	$\hat{r}_{\hat{R}_i}$
0	0.00	0.00	0.00
1	94,633.81	68,019.12	42,373.78
2	469,511.29	234,415.09	167,086.40
3	709,637.82	339,070.76	222,474.12
4	984,888.64	531,697.41	447,716.23
5	1,419,459.46	835,518.82	1,003,969.45
6	2,177,640.62	1,334,149.99	1,480,419.03
7	3,920,301.01	2,236,385.01	3,268,264.96
8	4,278,972.26	2,302,528.34	4,001,679.81
9	4,625,810.69	2,808,845.71	6,146,664.83
$\sum$	18,680,855.61	10,690,630.25	16,780,648.61

Table 6.4: Estimated reserves $\hat{\tilde{R}}_i$ for $i = 0, \ldots, I$.

6.3.2 Comparison with FCL Method

This section deals with the comparison of the examples of the FCL method as presented in Chapter 5 and the one proposed in this chapter. As a first obvious structural difference the appearance of asymmetric fuzzy development factors can be mentioned.

Comparing Tables 6.2 and 5.2 shows that the total spreads of the development factors in the FCL method is larger than in the AFCL one. For example, for the development factor $\hat{\tilde{f}}_1$ the total spread is 4.981 in the FCL method versus 2.003 in the AFCL one. Moreover, the support of the estimators in this chapter is not bounded below to one.

Likewise, this has an impact on the predictions of the ultimate claims. As these are derived by successively multiplying the last observation of an accident year with the estimated development factors the total spread of predictions of the ultimate claims is smaller for the AFCL method. As an example the prediction of the ultimate claim for accident year nine shall be considered for which the total spread is obviously the largest because the most multiplications need to be conducted. Compared to the FCL method the total spread of the ultimate claim for accident year nine can be reduced from over 24 million to approximately 9 million (cf. Tables 6.3 and 5.3).

The reserves are calculated as difference of the prediction of the ultimate claims and the last observations. Thus, the observation made above has an immediate influence on

accident-year i	$\beta = 0.1$	$\beta = 0.25$	$E_\beta(\hat{\hat{R}}_i)$ $\beta = 0.5$	$\beta = 0.75$	$\beta = 0.9$	chain-ladder reserve $\hat{R}_i$
0	0.00	0.00	0.00	0.00	0.00	0.00
1	66,143.90	74,423.37	88,222.48	102,021.59	110,301.06	94,633.81
2	372,378.82	402,491.43	452,679.12	502,866.80	532,979.42	469,511.29
3	568,179.68	610,295.55	680,488.66	750,681.77	792,797.63	709,637.82
4	768,010.62	841,466.64	963,893.34	1,086,320.05	1,159,776.07	984,888.64
5	1,093,674.46	1,231,636.08	1,461,572.12	1,691,508.15	1,829,469.77	1,419,459.46
6	1,651,294.08	1,862,386.75	2,214,207.88	2,566,029.01	2,777,121.69	2,177,640.62
7	3,077,341.00	3,490,189.75	4,178,271.00	4,866,352.25	5,279,200.99	3,920,301.01
8	3,442,918.50	3,915,734.11	4,703,760.13	5,491,786.15	5,964,601.76	4,278,972.26
9	3,669,163.37	4,340,826.66	5,460,265.47	6,579,704.29	7,251,367.58	4,625,810.69
$\sum$	14,709,104.43	16,769,450.35	20,203,360.20	23,637,270.06	25,697,615.98	18,680,855.61

Table 6.5: Expected values of the reserves $\hat{\hat{R}}_i$ $(i = 0, \ldots, I)$ for different choices of the "decision-maker risk parameter" β and corresponding chain-ladder reserves $\hat{R}_i$ $(i = 0, \ldots, I)$.

accident-year i	$\mathrm{Unc}_K(\hat{\hat{C}}_{i,J} \mid \mathcal{D}_I)$ $K = 1$
0	0.00
1	55,196.45
2	200,750.75
3	280,772.44
4	489,706.82
5	919,744.14
6	1,407,284.51
7	2,752,324.99
8	3,152,104.08
9	4,477,755.27
$\sum$	13,735,639.43

Table 6.6: Estimation uncertainty $\mathrm{Unc}_K(\hat{\hat{C}}_{i,J} \mid \mathcal{D}_I)$ for each accident year i, $i = 0, \ldots, I$, for a choice of $K = 1$.

the uncertainty of the predictions of the reserves which is reduced as well (cf. Tables 6.4 and 5.4). Especially the right spread of the aggregated reserve in the AFCL method is by approximately 28 million lower than in the FCL method. Hence, with the help of the AFCL method the actuary is able to describe a more precise picture of the situation he/she evaluates as appropriate.

This also becomes apparent in the defuzzified reserves, i.e. their expected values, shown in Tables 6.5 and 5.5. Whereas the expected value of the reserves is higher in the AFCL-model for a choice of $\beta = 0.1$ it is lower for all other choices presented. The differences are up to approximately 12 millions for a choice of $\beta = 0.9$. This means that if the actuary is interested in a conservative prediction of the reserves (shown by the high choice of β) the AFCL method indicates the actuary that already smaller reserves compared to the FCL are sufficient. Contrastingly, the FCL method would lead to higher reserves which could by costly for the insurer. This observation with the AFCL method also runs through to the estimated uncertainties given in Tables 6.6 and 5.6. For $K = 1$ the estimated uncertainties are higher in the FCL method for all accident years. It results in a difference for the uncertainty of aggregated ultimate claims of approximately 18 million.

6.4 Discussion

The model presented in this chapter shows another application of fuzzy methods to the CL reserving method. To our knowledge the implementation of fuzzy regression techniques is one of the rare publications which use linear optimization practices in claims reserving. In comparison to the FCL method the asymmetric coefficients allow to model a wide range of different claims development behaviors. It is possible to depict a more differentiated assessment of the actuary. If the data points diverge only either upward or downward (or with a strong trend in one direction) asymmetric spreads do not necessarily lead to a large total spread. This results in the advantage that the left spreads are not inevitably that large that the range of shown reserves has to go down to zero. Even though in the FCL method small values of the reserves are only assumed with small grades of membership this disadvantage can be circumvented. The AFCL method can hence depict a more precise (in the sense that the spreads of the reserves are more narrow) picture of the observed claims development behavior. In case of smaller spreads – which we assume the common situation in comparison to the FCL method – this reduces the uncertainty in the reserves. Of much greater significance is, however, the much lower reserve a conservative actuary (represented by a choice of a high parameter β) will set up according to the AFCL method in comparison to the FCL method without changing his or her attitude to uncertainty.

In practice the actuary sometimes deals with the problem that the estimated development factors do not represent the whole scope of what he or she thinks is realistic. Maybe, there are too less variations in the observations for a specific LoB. The presented method allows to react in these situations by choosing an additional positive "margin" α_j in the vector $\boldsymbol{\alpha}^*$, i.e. $\alpha_j > 0$. As a consequence, the spreads of the estimated AFCL-factors will increase and, thus, the spreads of the predicted reserves.

The usage of asymmetric coefficients is also a clear distinction from the models of Andrés Sánchez and Terceño Gómez (2003) and Kerkez (2013a,b). This also cushions the effect that the total spread is unnecessarily large. Furthermore, the introduced method provides an opportunity to defuzzify the reserves and measure the uncertainty.

Although TFNs are fitted here according to the procedure, the model is not restricted to them. The coefficients can be as well adapted to other representatives of the family of L-R FNs. If the data does not justify the utilization of triangular membership functions e.g. a Gaussian one (cf. Definition 2.19) can be also thinkable.

7 The Fuzzy Bornhuetter Ferguson Method

This chapter is based on a joint work with Jochen Heberle (cf. Heberle and Thomas 2016).[62]

Firstly, a motivation for the fuzzy Bornhuetter Ferguson method is given in Section 7.1. The model is described in Section 7.2. The chapter proceeds with a deduction of the claims reserve in Section 7.3 and a quantification of the prediction uncertainty for single and aggregated accident years in Section 7.4. An example of the presented method is given in Section 7.5 and the chapter concludes with a critical discussion in Section 7.6.

7.1 Motivation

The classical BF method is a purely computational claims reserving method. Purely computational means that there is no underlying stochastic framework. The method makes use of external information which can be given by expert knowledge, market statistics, organizational data, reinsurance data, etc. Depending on the source of information the data is likely to be vague or not. The more reliable the source is, the more one can assume that the data is not vague. On the contrary, if the data is influenced by subjective knowledge of an individual it might be assumed to be vague. In classical computational or stochastic models vagueness is not modeled at all.

[62] Parts of this paper were presented at the 18th International Congress on Insurance: Mathematics & Economics in Shanghai, China, 2014.

The BF method is a popular claims reserving method and is presented in detail in Section 4.4. A possibility to comprise vague information is given by the methodology of fuzzy sets (cf. Chapter 2). We follow a similar approach as in the fuzzy chain-ladder method (cf. Chapter 5 or Heberle and Thomas (2014)) and propose a method in which all a priori information is modeled by triangular fuzzy numbers (cf. Definition 2.18) in order to incorporate vagueness. Arithmetic operations with TFNs are easy to conduct and TFNs can still map a large part of the actuary's intuition of the uncertainty. Both the a priori estimates for the ultimate claims and the claims development pattern are modeled with TFNs. In the case that the parameters of the claims development pattern are estimated with the help of the inverse of the chain-ladder factors (cf. Equation (4.16)), one can also argue for the use of the theory of fuzzy sets as those development factors might be adjusted retrospectively due to subjective judgment.

In Chapter 5 TFNs are applied to the CL method in the way that the development factors are modeled as TFNs. Moreover, an estimator for the mode as well as for the left and right spread are proposed and an attempt is made to quantify the uncertainty. In contrast to the BF method the classical CL-method and the FCL model can not utilize a priori information.

The aim of this chapter is to take into consideration both a priori and vague information. Accordingly, the BF method is enhanced with fuzzy methods. It appears in the numerical example that the FBF method in contrast to the FCL method can lead to results in which the vagueness of the predicted reserve is considerably lower. Therefore, the method may give the actuary a better intuition about the predicted reserves if a priori information is available.

The starting point is the classical BF method as defined in Section 4.4. For the fuzzy part we use TFNs as defined in Definition 2.18 and the secant approximation for the fuzzy multiplication (cf. Equation (2.7)) and for the fuzzy inverse (cf. Equation (2.9)). As defuzzification method the concept of an expected value of a TFN as defined in Definition 2.38 is applied and as measure of fuzziness we use the uncertainty of a TFN as defined in Definition 2.50.

7.2 The Model

In this section the classical BF method is extended to a fuzzy Bornhuetter Ferguson method with the help of fuzzy set theory. In particular, FNs are applied, i.e. we assume that both the a priori information for the ultimate claims ν_i, $i \in \{0, \ldots, I\}$, and the parameters γ_j, $j \in \{0, \ldots, J\}$, are TFNs. Therefore, these can be written as $\tilde{\nu}_i = (\nu_i, l_{\nu_i}, r_{\nu_i})$ and $\tilde{\gamma}_j = (\gamma_j, l_{\gamma_j}, r_{\gamma_j})$, respectively. We aim to derive predictors for the ultimate claims $\tilde{C}_{i,J}$, $i \in \{1, \ldots, I\}$, for the outstanding loss liabilities for single accident years $\tilde{R}_i$, $i \in \{1, \ldots, I\}$ as well as for aggregated accident years $\tilde{R}$ and estimators for the uncertainty of the predicted ultimate claims for single accident years $\mathrm{Unc}_K\left(\hat{\tilde{C}}^{\mathrm{FBF}}_{i,J} \,\middle|\, \mathcal{D}_I\right)$, $i \in \{1, \ldots, I\}$ and for aggregated accident years $\mathrm{Unc}_K\left(\bigoplus_{i=1}^{I} \hat{\tilde{C}}^{\mathrm{FBF}}_{i,J} \,\middle|\, \mathcal{D}_I\right)$.

As a clarification, the FBF method models fuzziness and not randomness. As the a priori information often derives from subjective knowledge in form of expert knowledge fuzziness is also present in claims reserving.
Thus, we make the following model assumptions:

Model Assumptions 7.1 (Fuzzy Bornhuetter Ferguson (FBF) model)
For the fuzzy Bornhuetter Ferguson model we assume:

- There exist positive TFNs $\tilde{\nu}_i$, $i \in \{0, \ldots, I\}$, and also positive TFNs $\tilde{\gamma}_j$, $j \in \{0, \ldots, J\}$, such that

$$\tilde{C}_{i,0} = \tilde{\gamma}_0 \otimes \tilde{\nu}_i \tag{7.1}$$

$$\tilde{C}_{i,j+k} = \tilde{C}_{i,j} \oplus (\tilde{\gamma}_{j+k} \ominus \tilde{\gamma}_j) \otimes \tilde{\nu}_i \tag{7.2}$$

  holds true for all $i = 0, \ldots, I$, $j = 0, \ldots, J-1$ and $k = 1, \ldots, J-j$.

- The sums of incremental claims $\bigoplus_{i=0}^{I-j-1} \tilde{X}_{i,j+1}$ are non-negative with $\tilde{X}_{i,j+1} := \tilde{C}_{i,j+1} \ominus \tilde{C}_{i,j}$ for all $j \in \{0, \ldots, J-1\}$.

Remarks 7.2

- Since the real numbers can be embedded in the set of fuzzy numbers, i.e. every real number $a \in \mathbb{R}$ can be denoted as a TFN $\tilde{a} = (a, 0, 0)$, observable cumulative claims $(i + j \leq I)$ can be written as $\tilde{C}_{i,j} = (C_{i,j}, 0, 0)$.

- There is no need for an error term in Equations (7.1) and (7.2) as the uncertainty is included in the fuzzy numbers in this model.
- The sum of incremental claims is assumed to be non-negative due to technical reasons (cf. the choice of the following Estimator 7.3). However, individual incremental claims can be either negative or positive so that adjustments of the reserves in both directions are possible. This assumption is generally but not necessarily fulfilled in practice.

On the one hand the a priori information for the estimates of the ultimate claims $\tilde{\nu}_i$, $i \in \{0, \ldots, I\}$, and the estimates of the parameters of the claims development pattern can be both given externally, or on the other hand the parameters of the claims development pattern can be estimated with the help of development factors. The latter case leads to an estimator for the parameters $\tilde{\gamma}_j$, $j \in \{0, \ldots, J\}$ with $\tilde{\gamma}_J = (1, 0, 0)$ as given in the following Estimator 7.3. With the help of these estimators the ultimate claims $\tilde{C}_{i,J}$, $i \in \{1, \ldots, I\}$ can be predicted by

$$\hat{\tilde{C}}_{i,J} = C_{i,I-i} \oplus \left(\hat{\tilde{\gamma}}_J \ominus \hat{\tilde{\gamma}}_{I-i}\right) \otimes \hat{\tilde{\nu}}_i. \tag{7.3}$$

In the FBF method the parameters are assumed to be TFNs such that we need estimators for the mode as well as for the left and right spreads. The more vague the available data is, the wider spreads an actuary will choose. Likewise, he/she will opt for a narrower width in the opposite situation.

In practice actuaries tend to modify the estimated parameters in the claims development pattern due to subjective judgment. With the help of the FCL estimators as given in Estimator 5.5 we choose Estimator 7.3 as motivated by Equation (4.16). Hence, estimators for the mode and both spreads are given in Equation (7.4). In this case, the choice of the estimator does not necessarily lead to symmetric spreads.

Estimator 7.3 (FBF Estimator for the TFNs $\tilde{\gamma}_j$)
The TFNs $\tilde{\gamma}_j = (\gamma_j, l_{\gamma_j}, r_{\gamma_j})$, $j = 0, \ldots, J$, introduced in Model Assumptions 7.1 can be estimated by $\hat{\tilde{\gamma}}_j = (\hat{\gamma}_j, \hat{l}_{\hat{\gamma}_j}, \hat{r}_{\hat{\gamma}_j})$ with

$$\hat{\tilde{\gamma}}_j = \bigotimes_{k=j}^{J-1} \hat{\tilde{f}}_k^{-1}$$

for $j = 0, \dots, J-1$ *and*

$$\hat{\tilde{\gamma}}_J = (1, 0, 0). \tag{7.4}$$

Here, $\hat{\tilde{f}}_k = \left(\hat{f}_k, \hat{l}_{\hat{f}_k}, \hat{r}_{\hat{f}_k}\right)$ *is an estimator for the fuzzy chain-ladder factor* $\tilde{f}_k = (f_k, l_{f_k}, r_{f_k})$*,* $k \in \{0, \dots, J-1\}$*, and is given by*

$$\hat{f}_k = \frac{\sum\limits_{i=0}^{I-k-1} C_{i,k+1}}{\sum\limits_{i=0}^{I-k-1} C_{i,k}}$$

and

$$\hat{l}_{\hat{f}_k} = \hat{r}_{\hat{f}_k} = \frac{\sum\limits_{i=0}^{I-k-1} X_{i,k+1}}{\sum\limits_{i=0}^{I-k-1} C_{i,k}}$$

where $X_{i,k+1} = C_{i,k+1} - C_{i,k}$ *for* $i = 0, \dots, I$ *and* $k = 0, \dots, J-1$ *(cf. Estimator 5.5).*

Remarks 7.4

- The FCL factors $\tilde{f}_k = (f_k, l_{f_k}, r_{f_k})$, $k \in \{0, \dots, J-1\}$, have been introduced in Estimator 5.5 and in Heberle and Thomas (2014). Here, the sum of incremental claims are assumed to be non-negative (cf. Model Assumptions 7.1) due to technical reasons of the choice of the FCL-estimators.

- Generally, the FBF estimates $\hat{\tilde{\gamma}}_j$ are not symmetric (except for $\hat{\tilde{\gamma}}_J = (1, 0, 0)$), even though the FCL estimates $\hat{\tilde{f}}_k$ are so, i.e. $\hat{l}_{\hat{f}_k} = \hat{r}_{\hat{f}_k}$ holds true for all $k = 0, \dots, J-1$.

- Similarly to the FCL model the total spread of $\hat{\tilde{\gamma}}_j$ given by $\hat{l}_{\hat{\gamma}_j} + \hat{r}_{\hat{\gamma}_j}$ can also be interpreted as "uncertainty" of the FN $\tilde{\gamma}_j$.

- The FCL estimators $\hat{\tilde{f}}_k$, $k \in \{0, \dots, J-1\}$, are TFNs which are bounded below to 1 (cf. Remarks 5.6). Obviously, this does not hold true in the FBF model for $\hat{\tilde{\gamma}}_j$. However, this is not necessary and it can be shown that the left-border $\hat{\gamma}_j - \hat{l}_{\hat{\gamma}_j}$ of the support for $j \in \{0, \dots, J\}$ is not smaller than zero.
  We relax the conditions in the way that we assume that the support of the considered TFNs, i.e. of the FCL factors, is not bounded below to one but to zero. We show that the support of the estimators $\hat{\tilde{\gamma}}_j$ is bounded below to zero even in

this case. Therefore, let $\tilde{a} = (a, l_a, r_a)$ and $\tilde{b} = (b, l_b, r_b)$ be non-negative TFNs, i.e. $a - l_a \geq 0$ and $b - l_b \geq 0$. Then, we have for the product of the inverses of those two TFNs:

$$\begin{aligned}
\tilde{c} := (c, l_c, r_c) &:= \tilde{a}^{-1} \otimes \tilde{b}^{-1} \\
&= \left(\frac{1}{a}, \frac{r_a}{a(a+r_a)}, \frac{l_a}{a(a-l_a)}\right) \otimes \left(\frac{1}{b}, \frac{r_b}{b(b+r_b)}, \frac{l_b}{b(b-l_b)}\right) \\
&= \Big(\frac{1}{ab}, \frac{r_b}{ab(b+r_b)} + \frac{r_a}{ab(a+r_a)} - \frac{r_a r_b}{ab(a+r_a)(b+r_b)}, \\
&\qquad \frac{l_b}{ab(b-l_b)} + \frac{l_a}{ab(a-l_a)} + \frac{l_a l_b}{ab(a-l_a)(b-l_b)}\Big)
\end{aligned}$$

Here we get for the lower bound of the support

$$\begin{aligned}
c - l_c &= \frac{1}{ab} - \left(\frac{r_b}{ab(b+r_b)} + \frac{r_a}{ab(a+r_a)} - \frac{r_a r_b}{ab(a+r_a)(b+r_b)}\right) \\
&= \frac{1}{ab}\left(1 - \frac{r_b}{b+r_b} - \frac{r_a}{a+r_a} + \frac{r_a r_b}{(a+r_a)(b+r_b)}\right) \\
&= \frac{1}{ab}\,\frac{(a+r_a)(b+r_b) - r_b(a+r_a) - r_a(b+r_b) + r_a r_b}{(a+r_a)(b+r_b)} \\
&= \frac{1}{ab}\,\frac{ab}{(a+r_a)(b+r_b)} \\
&= \frac{1}{(a+r_a)(b+r_b)} > 0
\end{aligned}$$

As the support of the resulting product $\tilde{c}$ is also bounded below to zero we yield the proposition by induction.

- In the FCL model the expected values $E_{0.5}(\hat{\hat{\tilde{f}}}_j)$ for $j = 0, \ldots, J-1$ and a parameter $\beta = 0.5$ equal the classical CL estimators (cf. Remarks 5.6). Here, we do not yield an equivalent result for $\hat{\hat{\tilde{\gamma}}}_j$ as the fuzzy product and fuzzy inverse do not maintain the symmetric structure of the FCL estimators $\hat{\hat{\tilde{f}}}_j$.

7.3 Claims Reserves

The goal of the FBF method is to yield a predictor for the outstanding loss liabilities with the help of fuzzy methods. In Equation (7.3) a predictor for cumulative claims for accident years $i = 1, \ldots, I$ is given and used to derive a predictor for the outstanding loss liabilities for single accident years. The claims reserves $\tilde{R}_i$, $i \in \{1, \ldots, I\}$, are given by

$$\tilde{R}_i = \tilde{C}_{i,J} \ominus \tilde{C}_{i,I-i} \tag{7.5}$$

where $\tilde{C}_{i,I-i}$ is observable, i.e. $\tilde{C}_{i,I-i} = (C_{i,I-i}, 0, 0)$. At time $t = I$ only the observations $\mathcal{D}_I$ are observable so that the ultimate claims $\tilde{C}_{i,J}$, $i \in \{1, \ldots, I\}$, are unobservable. Hence, the claims reserves for single accident years need to be predicted. By substituting the quantities in Equation (7.5) by their estimates we yield a predictor for the claims reserves, i.e.

$$\hat{\tilde{R}}_i = \hat{\tilde{C}}_{i,J} \ominus \tilde{C}_{i,I-i}, \quad i = 1, \ldots, I. \tag{7.6}$$

For a given set of observations $\mathcal{D}_I$ the ultimate claims $\hat{\tilde{C}}_{i,J}$ for accident year $i = 1, \ldots, I$ can be presented as

$$\hat{\tilde{C}}_{i,J} = C_{i,I-i} \oplus \Big((1,0,0) \ominus \hat{\tilde{\gamma}}_{I-i} \Big) \otimes \hat{\tilde{\nu}}_i$$

(cf. Equation (7.3)). Thus, we can identify the predictor given in Equation (7.6) by

$$\begin{aligned}
\hat{\tilde{R}}_i &= \hat{\tilde{C}}_{i,J} \ominus \tilde{C}_{i,I-i} \\
&= C_{i,I-i} \oplus \Big((1,0,0) \ominus \hat{\tilde{\gamma}}_{I-i} \Big) \otimes \hat{\tilde{\nu}}_i \ominus \tilde{C}_{i,I-i} \\
&= \Big((1,0,0) \ominus \hat{\tilde{\gamma}}_{I-i} \Big) \otimes \hat{\tilde{\nu}}_i .
\end{aligned}$$

We remark that since at time $t = I$ the cumulative claim $\tilde{C}_{i,I-i}$ is observable and therefore crisp, we do not yield a fuzzy zero but a crisp zero in the last above equation.

7.4 Prediction Uncertainty

Actuaries are not only interested in the prediction of the outstanding loss liabilities but also in the prediction's accuracy. Therefore, we derive an estimator for the ultimate claim uncertainty in Section 7.4.1 as well as for aggregated accident years in Section 7.4.2.

7.4.1 Single Accident Years

In classical reserving methods the prediction uncertainty is often measured with the mean square error of prediction (MSEP) – an approach which cannot be applied here since we are dealing with fuzzy methods. Hence, we use a similar method as in Section 5.4.1 which makes use of Definition 2.50. An estimator for the ultimate claim uncertainty is yielded as given in Estimator 7.5.

Estimator 7.5 (Ultimate claim uncertainty)
Given the observations $\mathcal{D}_I$ and a scalar $K \in \mathbb{R}^+$, the uncertainty of the ultimate claim $\hat{\tilde{C}}_{i,J}$ for accident year $i \in \{1, \ldots, I\}$ in the FBF model can be represented as

$$\mathrm{Unc}_K\left(\hat{\tilde{C}}_{i,J}^{FBF} \,\middle|\, \mathcal{D}_I\right) = \tfrac{1}{2}K\Big((1-\hat{\gamma}_{I-i})\left(\hat{l}_{\hat{\nu}_i} + \hat{r}_{\hat{\nu}_i}\right) + \left(\hat{l}_{\hat{\gamma}_{I-i}} + \hat{r}_{\hat{\gamma}_{I-i}}\right)\hat{\nu}_i \\ + \hat{l}_{\hat{\gamma}_{I-i}}\hat{r}_{\hat{\nu}_i} - \hat{r}_{\hat{\gamma}_{I-i}}\hat{l}_{\hat{\nu}_i}\Big).$$

Proof. Assuming that the observations $\mathcal{D}_I$ are given we can state the ultimate claims $\hat{\tilde{C}}_{i,J}$ for accident years $i \in \{1, \ldots, I\}$ in the following way:

$$\begin{aligned}
\hat{\tilde{C}}_{i,J} &= \tilde{C}_{i,I-i} \oplus \left((1,0,0) \ominus \hat{\tilde{\gamma}}_{I-i}\right) \otimes \hat{\tilde{\nu}}_i \\
&= (C_{i,I-i},0,0) \oplus \left((1,0,0) \ominus \left(\hat{\gamma}_{I-i}, \hat{l}_{\hat{\gamma}_{I-i}}, \hat{r}_{\hat{\gamma}_{I-i}}\right)\right) \otimes \left(\hat{\nu}_i, \hat{l}_{\hat{\nu}_i}, \hat{r}_{\hat{\nu}_i}\right) \\
&= (C_{i,I-i},0,0) \oplus \left(1-\hat{\gamma}_{I-i}, \hat{r}_{\hat{\gamma}_{I-i}}, \hat{l}_{\hat{\gamma}_{I-i}}\right) \otimes \left(\hat{\nu}_i, \hat{l}_{\hat{\nu}_i}, \hat{r}_{\hat{\nu}_i}\right) \\
&= (C_{i,I-i},0,0) \oplus \Big((1-\hat{\gamma}_{I-i})\hat{\nu}_i, (1-\hat{\gamma}_{I-i})\hat{l}_{\hat{\nu}_i} + \hat{\nu}_i\hat{r}_{\hat{\gamma}_{I-i}} - \hat{r}_{\hat{\gamma}_{I-i}}\hat{l}_{\hat{\nu}_i}, \\
&\qquad (1-\hat{\gamma}_{I-i})\hat{r}_{\hat{\nu}_i} + \hat{\nu}_i\hat{l}_{\hat{\gamma}_{I-i}} + \hat{l}_{\hat{\gamma}_{I-i}}\hat{r}_{\hat{\nu}_i}\Big) \\
&= \Big(C_{i,I-i} + (1-\hat{\gamma}_{I-i})\,\hat{\nu}_i, (1-\hat{\gamma}_{I-i})\,\hat{l}_{\hat{\nu}_i} + \hat{r}_{\hat{\gamma}_{I-i}}\hat{\nu}_i - \hat{r}_{\hat{\gamma}_{I-i}}\hat{l}_{\hat{\nu}_i}, \\
&\qquad (1-\hat{\gamma}_{I-i})\,\hat{r}_{\hat{\nu}_i} + \hat{l}_{\hat{\gamma}_{I-i}}\hat{\nu}_i + \hat{l}_{\hat{\gamma}_{I-i}}\hat{r}_{\hat{\nu}_i}\Big)
\end{aligned} \tag{7.7}$$

We can derive the uncertainty $\mathrm{Unc}_K\left(\hat{\tilde{C}}_{i,J}^{\mathrm{FBF}} \,\middle|\, \mathcal{D}_I\right)$ by applying Definition 2.50 to the representation of the ultimate claim shown in Equation (7.7).

$$\begin{aligned}
\mathrm{Unc}_K\left(\hat{\tilde{C}}_{i,J}^{\mathrm{FBF}} \,\middle|\, \mathcal{D}_I\right) &= \tfrac{1}{2}K\Big((1-\hat{\gamma}_{I-i})\,\hat{l}_{\hat{\nu}_i} + \hat{r}_{\hat{\gamma}_{I-i}}\hat{\nu}_i - \hat{r}_{\hat{\gamma}_{I-i}}\hat{l}_{\hat{\nu}_i} + (1-\hat{\gamma}_{I-i})\,\hat{r}_{\hat{\nu}_i} \\
&\qquad + \hat{l}_{\hat{\gamma}_{I-i}}\hat{\nu}_i + \hat{l}_{\hat{\gamma}_{I-i}}\hat{r}_{\hat{\nu}_i}\Big) \\
&= \tfrac{1}{2}K\Big((1-\hat{\gamma}_{I-i})\left(\hat{l}_{\hat{\nu}_i} + \hat{r}_{\hat{\nu}_i}\right) + \left(\hat{l}_{\hat{\gamma}_{I-i}} + \hat{r}_{\hat{\gamma}_{I-i}}\right)\hat{\nu}_i \\
&\qquad + \hat{l}_{\hat{\gamma}_{I-i}}\hat{r}_{\hat{\nu}_i} - \hat{r}_{\hat{\gamma}_{I-i}}\hat{l}_{\hat{\nu}_i}\Big)
\end{aligned}$$

□

With help of Estimator 7.5 a representation for the ultimate claim uncertainty for aggregated accident years is deduced in the following section.

7.4.2 Aggregated Accident Years

The concern does not only lie in the prediction uncertainty for single accident years but also in the aggregated ones. Therefore, we derive an estimator for the prediction uncertainty of the aggregated ultimate claim $\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J}$ given the observations $\mathcal{D}_I$. The aggregated ultimate claim can be perceived as the sum of the individual ultimate claims, i.e.

$$\begin{aligned}
\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J} &= \bigoplus_{i=1}^{I} \Big((C_{i,I-i}, 0, 0) \oplus \Big((1,0,0) \ominus \hat{\hat{\gamma}}_{I-i} \Big) \otimes \hat{\hat{\nu}}_i \Big) \\
&= \bigoplus_{i=1}^{I} \Big(C_{i,I-i} + (1 - \hat{\gamma}_{I-i})\, \hat{\nu}_i, (1 - \hat{\gamma}_{I-i})\, \hat{l}_{\hat{\nu}_i} + \hat{r}_{\hat{\gamma}_{I-i}} \hat{\nu}_i - \hat{r}_{\hat{\gamma}_{I-i}} \hat{l}_{\hat{\nu}_i}, \\
&\qquad (1 - \hat{\gamma}_{I-i})\, \hat{r}_{\hat{\nu}_i} + \hat{l}_{\hat{\gamma}_{I-i}} \hat{\nu}_i + \hat{l}_{\hat{\gamma}_{I-i}} \hat{r}_{\hat{\nu}_i} \Big),
\end{aligned} \tag{7.8}$$

That is why we can immediately state an estimator for the uncertainty of the aggregated ultimate claims as given in Estimator 7.6.

Estimator 7.6 (Uncertainty of the aggregated ultimate claims)
Given the observations $\mathcal{D}_I$ and a parameter $K \in \mathbb{R}^+$, the uncertainty of the aggregated ultimate claims $\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J}$ in the FBF model can be represented as

$$\mathrm{Unc}_K \left(\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J}^{FBF} \,\middle|\, \mathcal{D}_I \right) = \sum_{i=1}^{I} \mathrm{Unc}_K \left(\hat{\hat{C}}_{i,J}^{FBF} \,\middle|\, \mathcal{D}_I \right). \tag{7.9}$$

Proof. Based on Equation (7.8) we compute the uncertainty with the help of Definition 2.50 and Estimator 7.5:

$$\begin{aligned}
\mathrm{Unc}_K \left(\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J}^{\mathrm{FBF}} \,\middle|\, \mathcal{D}_I \right) &= \mathrm{Unc}_K \Big(\bigoplus_{i=1}^{I} \Big(C_{i,I-i} + (1 - \hat{\gamma}_{I-i})\, \hat{\nu}_i, \\
&\qquad (1 - \hat{\gamma}_{I-i})\, \hat{l}_{\hat{\nu}_i} + \hat{r}_{\hat{\gamma}_{I-i}} \hat{\nu}_i - \hat{r}_{\hat{\gamma}_{I-i}} \hat{l}_{\hat{\nu}_i}, \\
&\qquad (1 - \hat{\gamma}_{I-i})\, \hat{r}_{\hat{\nu}_i} + \hat{l}_{\hat{\gamma}_{I-i}} \hat{\nu}_i + \hat{l}_{\hat{\gamma}_{I-i}} \hat{r}_{\hat{\nu}_i} \Big) \Big) \\
&= \mathrm{Unc}_K \Big(\sum_{i=1}^{I} \left(C_{i,I-i} + (1 - \hat{\gamma}_{I-i})\, \hat{\nu}_i \right),
\end{aligned}$$

$$\sum_{i=1}^{I}\left(\left(1-\hat{\gamma}_{I-i}\right)\hat{l}_{\hat{\nu}_i}+\hat{r}_{\hat{\gamma}_{I-i}}\hat{\nu}_i-\hat{r}_{\hat{\gamma}_{I-i}}\hat{l}_{\hat{\nu}_i}\right),$$

$$\sum_{i=1}^{I}\left(\left(1-\hat{\gamma}_{I-i}\right)\hat{r}_{\hat{\nu}_i}+\hat{l}_{\hat{\gamma}_{I-i}}\hat{\nu}_i+\hat{l}_{\hat{\gamma}_{I-i}}\hat{r}_{\hat{\nu}_i}\right)\Big)$$

$$= \sum_{i=1}^{I}\tfrac{1}{2}K\Big(\left(1-\hat{\gamma}_{I-i}\right)\left(\hat{l}_{\hat{\nu}_i}+\hat{r}_{\hat{\nu}_i}\right)+\left(\hat{l}_{\hat{\gamma}_{I-i}}+\hat{r}_{\hat{\gamma}_{I-i}}\right)\hat{\nu}_i$$

$$+\hat{l}_{\hat{\gamma}_{I-i}}\hat{r}_{\hat{\nu}_i}-\hat{r}_{\hat{\gamma}_{I-i}}\hat{l}_{\hat{\nu}_i}\Big)$$

$$= \sum_{i=1}^{I}\mathrm{Unc}_K\left(\hat{\tilde{C}}_{i,J}^{\mathrm{FBF}}\,\middle|\,\mathcal{D}_I\right)$$

$\square$

Remark 7.7

In comparison to the MSEP for aggregated accident years in classical reserving methods there are no covariance terms in Equation (7.9). In fact, the uncertainties of different accident years cannot offset each other. One opportunity to model dependencies among different accident years as well would be to go over to the theory of fuzzy random variables.

7.5 Example

In the following we apply the FBF method to a given run-off triangle. We consider the *paid* run-off triangle as given in Dahms (2008, p. 15) which is shown in Table 7.1. As a priori estimates we use the last observed diagonal in the *incurred* run-off triangle also given in Dahms (2008, p. 15). This last diagonal can be found again in the first column in Table 7.2. We like to keep in mind that the example presented in Dahms (2008) regards randomness whereas fuzziness is addressed here. Thus, the results can not be compared.

The FBF method asks for fuzzy a priori estimates, that is why we assume that the given information consists of fuzzy and not crisp values. In this way, vagueness is added. The a priori estimates are given in Table 7.2.

The a priori information often originates from expert knowledge and the spreads need to be chosen by the actuary. The actuary will choose the spreads under different considerations: On the one hand vagueness rises the more values for a given accident year need to be predicted (cf. Model Assumptions 7.1 and Estimator 7.3), on the other

accident year i	development year j 0	1	2	3	4	5	6	7	8	9
0	1,216,632	1,347,072	1,786,877	2,281,606	2,656,224	2,909,307	3,283,388	3,587,549	3,754,403	3,921,258
1	798,924	1,051,912	1,215,785	1,349,939	1,655,312	1,926,210	2,132,833	2,287,311	2,567,056	
2	1,115,636	1,387,387	1,930,867	2,177,002	2,513,171	2,931,930	3,047,368	3,182,511		
3	1,052,161	1,321,206	1,700,132	1,971,303	2,298,349	2,645,113	3,003,425			
4	808,864	1,029,523	1,229,626	1,590,338	1,842,662	2,150,351				
5	1,016,862	1,251,420	1,698,052	2,105,143	2,385,339					
6	948,312	1,108,791	1,315,524	1,487,577						
7	917,530	1,082,426	1,484,405							
8	1,001,238	1,376,124								
9	841,930									

Table 7.1: Observed cumulative claims payments $C_{i,j}$.

i	ν_i	l_{ν_i}	r_{ν_i}
0	3,921,258	0	0
1	2,919,955	100,000	100,000
2	3,257,827	200,000	200,000
3	3,413,921	300,000	300,000
4	3,298,998	400,000	400,000
5	3,702,427	500,000	500,000
6	3,704,113	600,000	600,000
7	4,408,097	700,000	700,000
8	4,132,757	800,000	800,000
9	3,045,376	900,000	900,000

Table 7.2: Given a priori information $\tilde{\nu}_i = (\nu_i, l_{\nu_i}, r_{\nu_i})$.

hand large absolute values can accommodate a higher vagueness. The left and right spreads of the a priori information given in Table 7.2 are increasing for later accident years $i \in \{1, \ldots, 9\}$. The way of assigning a membership function to the given a priori estimates is just an example. It is driven by the fact that vagueness rises the more values need to be predicted. For simplicity's sake we assume the membership functions to be symmetric. For accident year zero ($i = 0$) no ultimate claim needs to be predicted so that we assign no spread, i.e. for $i = 0$ the a priori information is a crisp number.

In Table 7.3 the estimated FCL factors $\hat{\tilde{f}}_j = \left(\hat{f}_j, \hat{l}_{\hat{f}_j}, \hat{r}_{\hat{f}_j}\right)$, $j \in \{0, \ldots, J-1\}$ as well as the estimated parameters $\hat{\tilde{\gamma}} = \left(\hat{\gamma}_j, \hat{l}_{\hat{\gamma}_j}, \hat{r}_{\hat{\gamma}_j}\right)$, $j \in \{0, \ldots, J\}$, of the FBF method are given. Both are written down in three lines, in the first line the mode and in the second and third line, the left and right spread, respectively, are given.

	development year j									
	0	1	2	3	4	5	6	7	8	9
$\hat{f}_j$	1.2343	1.2904	1.1918	1.1635	1.1457	1.1013	1.0702	1.0760	1.0444	–
$\hat{l}_{\hat{f}_j}$	0.2343	0.2904	0.1918	0.1635	0.1457	0.1013	0.0702	0.0760	0.0444	–
$\hat{r}_{\hat{f}_j}$	0.2343	0.2904	0.1918	0.1635	0.1457	0.1013	0.0702	0.0760	0.0444	–
$\hat{\gamma}_j$	0.2984	0.3683	0.4753	0.5664	0.6590	0.7550	0.8315	0.8898	0.9574	1
$\hat{l}_{\hat{\gamma}_j}$	0.1928	0.2132	0.2301	0.2272	0.2088	0.1737	0.1324	0.0926	0.0391	0
$\hat{r}_{\hat{\gamma}_j}$	0.7016	0.6317	0.5247	0.4336	0.3410	0.2450	0.1685	0.1102	0.0426	0

Table 7.3: Estimated FCL factors $\hat{\tilde{f}}_j$ and parameters $\hat{\tilde{\gamma}}_j$ of the FBF method for $j = 0, \ldots, J-1$. The parameter for $j = J$ is given by $\hat{\tilde{\gamma}}_J = (1, 0, 0)$.

By means of the calculated factors the run-off triangle can be filled up using Equation (7.2). The development triangle with predicted lower right part is shown in Table 7.4.

In Table 7.5 the predicted FBF reserves $\hat{\tilde{R}}_i^{\text{FBF}}$, $i \in \{0, \ldots, I\}$, as well as the aggregated FBF reserve $\hat{\tilde{R}}^{\text{FBF}}$ are presented on the right hand side. On the left side, the FCL reserves $\hat{\tilde{R}}_i^{\text{FCL}}$, $i \in \{0, \ldots, I\}$, and the aggregated FCL reserve $\hat{\tilde{R}}^{\text{FCL}}$ are shown. In the first column for each procedure the mode of the fuzzy reserve is given and in the second and third column the left and right spread, respectively, are written down. The support of the fuzzy reserves ranges from $\hat{R}_i - \hat{l}_{\hat{R}_i}$ to $\hat{R}_i + \hat{r}_{\hat{R}_i}$. Thus, reserves for both methods can amount down to zero as the difference of the mode and the left spread yields zero for all accident years. However, small values for the reserve only have a little grade of membership as the slope of the left part of the membership function is low. The modes of the FBF reserves are higher except for accident years two and three. The same can be observed for the left spreads. The right spreads of the FBF reserves are much narrower with an exception for accident year one. The total right spread of the aggregated FBF reserve is only about 45% of the total right spread of the FCL reserve. This is due to the underlying arithmetic operations: In contrast to the FBF method we yield the FCL reserve by various fuzzy multiplications which lead to larger spreads.

The expected values of the FBF reserves $\hat{\tilde{R}}_i^{\text{FBF}}$, $i \in \{0, \ldots, I\}$, for different values of the decision-maker risk parameter and the uncertainty of these predictions are presented in Table 7.6. As one might expect, the expected aggregated reserve is higher, the more conservative the actuary is in the sense of a high parameter β. The reason is that the expected value puts more weight on the right spread, the higher the parameter

	development year j									
$\hat{\hat{C}}_{i,j}$	0	1	2	3	4	5	6	7	8	9
$\hat{C}_{0,j}$	1,216,632	1,347,072	1,786,877	2,281,606	2,656,224	2,909,307	3,283,388	3,587,549	3,754,403	3,921,258
$\hat{l}_{\hat{C}_{0,j}}$	0	0	0	0	0	0	0	0	0	0
$\hat{r}_{\hat{C}_{0,j}}$	0	0	0	0	0	0	0	0	0	0
$\hat{C}_{1,j}$	798,924	1,051,912	1,215,785	1,349,939	1,655,312	1,926,210	2,132,833	2,287,311	2,567,056	2,691,304
$\hat{l}_{\hat{C}_{1,j}}$	0	0	0	0	0	0	0	0	0	124,248
$\hat{r}_{\hat{C}_{1,j}}$	0	0	0	0	0	0	0	0	0	122,268
$\hat{C}_{2,j}$	1,115,636	1,387,387	1,930,867	2,177,002	2,513,171	2,931,930	3,047,368	3,182,511	3,402,877	3,541,502
$\hat{l}_{\hat{C}_{2,j}}$	0	0	0	0	0	0	0	0	469,974	358,990
$\hat{r}_{\hat{C}_{2,j}}$	0	0	0	0	0	0	0	0	480,982	342,357
$\hat{C}_{3,j}$	1,052,161	1,321,206	1,700,132	1,971,303	2,298,349	2,645,113	3,003,425	3,202,572	3,433,496	3,578,763
$\hat{l}_{\hat{C}_{3,j}}$	0	0	0	0	0	0	0	830,740	684,258	575,338
$\hat{r}_{\hat{C}_{3,j}}$	0	0	0	0	0	0	0	918,446	687,522	542,255
$\hat{C}_{4,j}$	808,864	1,029,523	1,229,626	1,590,338	1,842,662	2,150,351	2,402,587	2,595,030	2,818,181	2,958,558
$\hat{l}_{\hat{C}_{4,j}}$	0	0	0	0	0	0	1,124,603	1,032,680	904,473	808,206
$\hat{r}_{\hat{C}_{4,j}}$	0	0	0	0	0	0	1,296,392	1,103,948	880,798	740,421
$\hat{C}_{5,j}$	1,016,862	1,251,420	1,698,052	2,105,143	2,385,339	2,740,732	3,023,814	3,239,791	3,490,230	3,647,773
$\hat{l}_{\hat{C}_{5,j}}$	0	0	0	0	0	1,696,126	1,602,149	1,503,996	1,366,302	1,262,434
$\hat{r}_{\hat{C}_{5,j}}$	0	0	0	0	0	1,955,127	1,672,045	1,456,068	1,205,628	1,048,085
$\hat{C}_{6,j}$	948,312	1,108,791	1,315,524	1,487,577	1,830,534	2,186,089	2,469,300	2,685,375	2,935,928	3,093,543
$\hat{l}_{\hat{C}_{6,j}}$	0	0	0	0	2,049,617	1,998,085	1,915,813	1,827,402	1,701,737	1,605,966
$\hat{r}_{\hat{C}_{6,j}}$	0	0	0	0	2,500,861	2,145,306	1,862,096	1,646,020	1,395,467	1,237,852
$\hat{C}_{7,j}$	917,530	1,082,426	1,484,405	1,886,218	2,294,356	2,717,486	3,054,522	3,311,663	3,609,835	3,797,406
$\hat{l}_{\hat{C}_{7,j}}$	0	0	0	2,851,831	2,848,685	2,785,518	2,685,958	2,579,367	2,428,119	2,313,000
$\hat{r}_{\hat{C}_{7,j}}$	0	0	0	3,453,681	3,045,544	2,622,414	2,285,378	2,028,236	1,730,064	1,542,493
$\hat{C}_{8,j}$	1,001,238	1,376,124	1,818,115	2,194,830	2,577,475	2,974,175	3,290,159	3,531,238	3,810,786	3,986,641
$\hat{l}_{\hat{C}_{8,j}}$	0	0	2,957,493	3,020,729	3,033,720	2,993,348	2,916,928	2,831,094	2,706,712	2,610,516
$\hat{r}_{\hat{C}_{8,j}}$	0	0	3,725,537	3,348,822	2,966,178	2,569,478	2,253,493	2,012,413	1,732,866	1,557,011
$\hat{C}_{9,j}$	841,930	1,054,861	1,380,559	1,658,155	1,940,121	2,232,444	2,465,289	2,642,937	2,848,932	2,978,517
$\hat{l}_{\hat{C}_{9,j}}$	0	2,025,490	2,157,918	2,233,720	2,277,732	2,288,701	2,268,946	2,236,152	2,182,128	2,136,587
$\hat{r}_{\hat{C}_{9,j}}$	0	3,315,686	2,989,988	2,712,392	2,430,426	2,138,103	1,905,258	1,727,609	1,521,614	1,392,029

Table 7.4: Filled run-off triangle with observed cumulative claims $C_{i,j}$ $(i+j \leq I)$ and predicted cumulative claims $\hat{\hat{C}}_{i,j}$ $(i+j > I)$.

accident year i	FCL reserves $\hat{\hat{R}}_i^{\text{FCL}}$ $\hat{R}_i$	$\hat{l}_{\hat{R}_i}$	$\hat{r}_{\hat{R}_i}$	FBF reserves $\hat{\hat{R}}_i^{\text{FBF}}$ $\hat{R}_i$	$\hat{l}_{\hat{R}_i}$	$\hat{r}_{\hat{R}_i}$
0	0.0	0.0	0.0	0.0	0.0	0.0
1	114,086.3	114,086.3	114,086.3	124,248.2	124,248.2	122,268.8
2	394,120.9	394,120.9	415,624.9	358,990.7	358,990.7	342,357.1
3	608,749.5	608,749.5	684,079.9	575,338.4	575,338.4	542,255.1
4	697,741.6	697,741.6	850,872.8	808,206.9	808,206.9	740,421.2
5	1,234,156.7	1,234,156.7	1,678,973.3	1,262,434.2	1,262,434.2	1,048,085.3
6	1,138,623.3	1,138,623.3	1,758,326.0	1,605,966.4	1,605,966.4	1,237,852.2
7	1,638,793.4	1,638,793.4	2,930,186.1	2,313,000.8	2,313,000.8	1,542,493.7
8	2,359,938.9	2,359,938.9	5,134,598.4	2,610,516.5	2,610,516.5	1,557,011.7
9	1,979,400.9	1,979,400.9	5,149,050.9	2,136,587.4	2,136,587.4	1,392,029.4
$\sum$	10,165,611.6	10,165,611.6	18,715,798.7	11,795,289.5	11,795,289.5	8,524,774.5

Table 7.5: Predicted FCL and FBF reserves for individual accident years $i \in \{0, \ldots, I\}$ and for aggregated accident years.

parameter	FCL	FBF
β	$E_\beta(\hat{\hat{R}}^{\text{FCL}})$	$E_\beta(\hat{\hat{R}}^{\text{FBF}})$
0.1	6,526,876	6,913,648
0.25	8,692,982	8,437,653
0.5	12,303,158	10,977,661
0.75	15,913,335	13,517,669
0.9	18,079,440	15,041,674
K	$\text{Unc}_K\left(\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J}^{\text{FCL}} \,\middle|\, \mathcal{D}_I\right)$	$\text{Unc}_K\left(\bigoplus_{i=1}^{I} \hat{\hat{C}}_{i,J}^{\text{FBF}} \,\middle|\, \mathcal{D}_I\right)$
1	14,440,705	10,160,032

Table 7.6: Expected reserves for different choices of the “decision-maker risk parameter” $\beta \in [0, 1]$ and the uncertainty of the predicted aggregated ultimate claims.

β. Furthermore, beginning at least at $\beta = 0.25$ the FBF method leads to smaller expected reserves (than with the FCL method) for this example. Apparently, the total spread of the aggregated reserve for the FBF method is narrower and, therefore, the corresponding uncertainty is lower (see Table 7.6). Hence, the presented method leads to more stable results in this example. This allows the conservative actuary (choice of a higher β) to choose a more optimistic prediction of the reserves than with the FCL method.

7.6 Discussion

Fuzzy set theory offers means to model uncertainty, but no stochastic randomness. Since a priori information can originate from expert knowledge, market statistics, organizational data, etc. they can be tainted with vagueness. The presented FBF method does not only provide an opportunity to model the a priori estimates for the ultimate claims, but also to model uncertainties with the parameters in the claims development pattern.

Here, we take a similar approach as for the FCL method (cf. Chapter 5) which also offers a possibility to model uncertainty in a claims reserving context. In this connection the following advantages of the FBF method over the FCL model can be highlighted. If a priori information is available, it should be used and the presented FBF method offers an opportunity to implement it. Moreover, the predictions are more precise due to the underlying arithmetic operations. A disadvantage of the FCL method is the rapidly rising prediction uncertainty for later accident years. This drawback can be overcome here as the presented method uses more fuzzy sums and less fuzzy multiplications which leads to a slighter increase in the spreads.

This chapter deals with modeling a priori information. However, a different approach as e.g. with the Bayes approach is conducted here. An advantage lies in the fact that the FBF method is easy to implement and e.g. no a priori distribution is needed.

In contrast to the FCL method the symmetry of the parameters cannot be maintained. The presented methods only deal with the modeling of uncertainty. In order to also map stochastic randomness one needs to go over to fuzzy random variables (cf. Section 2.4). Then, both sources of uncertainty, vagueness and randomness, can be implemented.

Moreover, dependencies within different accident or calendar years could be considered. However, the presented FBF method offers a first step in modeling vagueness in a BF reserving context.

In this method TFNs are utilized as the computational effort is low and they are easy to interpret. Of course, there might be situations in which other shapes of membership functions are more suitable as e.g. Gaussian or exponential ones (cf. Definitions 2.19 and 2.21).

8 | Conclusion

Estimating outstanding loss liabilities is one of the crucial tasks of an actuary in a non-life insurance company. As it amounts to a huge position on the liabilities side of the balance sheet of an insurance company neither overestimating nor underestimating is reasonable. If the reserve is too low, claims might not be paid. Moreover, reserves serve as a basis for pricing calculations. In the case of low premiums due to the reserve calculations, it might result in non-profitable business. Concerning too high reserves, this can lead to a lower stated profit and, thus, the equity base is smaller. Taking into account pricing considerations, higher reserves can cause higher premiums, especially in long-tail LoBs, and, therefore, maybe result in less new business for the insurance company. Therefore, over the last decades a number of stochastic methods in the field of claims reserving have been evolved. Beyond that, methods incorporating credibility and Bayes theory, Bootstrapping methods, Kalman filter, etc. exist. These methods – except for the purely computational ones – take into account stochastic randomness. What has been left out in these considerations is the uncertainty due to vagueness which can arise e.g. due to subjective judgment of the reserving actuary. A means to model vagueness is given by fuzzy set theory. The object of this work is the presentation of applications in claims reserving modeling vagueness. A systematic approach with the help of fuzzy methods is shown to model intuitive decisions by an actuary or methods involving expert knowledge. In the latter case not only a crisp specification but a range of possible values can be depicted.

As a first step a general introduction to the topic is given (cf. Chapter 1). Motives and reasons for the utilizations are mentioned and possible gaps in the current state of literature are pointed out. Fuzzy methods serve as a means to model vagueness which can comprise e.g. subjective judgment, linguistic inaccuracies and data tainted with errors. The necessary definitions and concepts are provided in Chapter 2.

The present work provides an overview of publications of fuzzy methods in actuarial science (cf. Chapter 3 and Section 4.6). A special emphasis is put on applications in the field of claims reserving. With this particular survey as well as a discussion of the advantages and disadvantages of those applications – to our best knowledge – a first structured review is accomplished.

Proceeding from the observation that development factors derived with the CL method might be altered at a later stage by an actuary the FCL model in Chapter 5 is set up. When comparing with earlier publications in this field, it distinguishes by its use of fuzzy numbers and arithmetic instead of fuzzy regression. Since claims reserving methods are not all regression models by nature this is an obvious enhancement of the current literature. In accordance to the classical CL-model assumptions are made. In comparison to the stochastic framework, estimators for the mode and both spreads are needed. This is directly related to the prediction of the ultimate claims which are deduced from successively fuzzy multiplying the last observation with the estimated fuzzy development factors. Subsequently, predictions for the outstanding loss liabilities are derived and defuzzification methods serve as a means to determine a crisp reserve. A real number is required since the reserve is a figure in the balance sheet. Regulatory frameworks ask for a crisp determination and not the specification of a range of values. Unlike previous articles concerning fuzzy claims reserving the prediction uncertainty is quantified.

Taking up an idea from the current literature a fuzzy regression model is used for predicting reserves in the context of the CL-model (cf. Chapter 6). Its representation as a weighted multiple linear regression model is used to apply the fuzzy regression model by Ishibuchi and Nii (2001). A regression "tube" is derived which contains all data points. The width of the tube can be adjusted subjectively by the assessment of the actuary. Reasons for this can be the validation of the tube in the sense that the "usual" (in the view of the actuary) scope of values is covered or the reliability of the data. Besides the statement of the fuzzy reserving method, predictions for the ultimate claim are derived. As this procedure as well as the FCL model utilize the CL method the results of an numerical example are critically compared with regard to the predictions of the ultimate claims and reserves as well as of the prediction uncertainty.

The third presented approach ties in with the classical BF method (cf. Section 4.4). The classical BF method is especially useful if a priori information for the ultimate claims' estimates is available. Often this information arises from expert knowledge. The FBF method allows for the application of BF even if the a priori estimates are liable to subjective judgment. Hence, it combines the advantages of the classical BF method with fuzzy set theory. Following the classical approach model assumptions are being made. Consequently, estimators for the development factors are chosen in order to derive predictions for the ultimate claims and, thus, the reserves. In accordance with the methods in previous chapters the prediction uncertainty is quantified. The consideration finishes with a critical discussion of the introduced method.

As a result of the present work fuzzy set theory can be seen as a means to encounter the question how to model vagueness due to subjective judgment in the field of claims reserving. Various approaches from the current state of literature as well as new ones are discussed. Still, there is room for further research and improvements.

First of all, TFNs are applied in the methods described in Chapters 5 and 7 as well as fuzzy coefficients in the model presented in Chapter 6. As elaborated on before, the use of differently shaped membership functions as e.g. Gaussian, exponential, quasi-exponential, could be a suitable extension depending on the situation. For example in the case of a highly reliable data source, an actuary opts for an exponential membership function such that values close to the calculated development factor possess a high grade of membership and then the function decreases rapidly for the left and right spread, respectively. The utilization of other members of the family of L-R FNs is to the detriment of the manageability of the model. That is the reason why TFNs are applied which can be also interpreted intuitively.

The models shown in Chapters 5, 6 and 7 all have in common that they are univariate. A further field for research is the generalization to the multivariate case, i.e. the simultaneous consideration of several correlated run-off portfolios. This can be motivated by the fact that multiple LoB are taken into account or that a portfolio needs to be subdivided into smaller portfolios for homogeneity assumptions in order to reasonably apply reserving methods like e.g. the multivariate CL-model.

From a global point of view, more fields of applications in actuarial science can be identified in which one can come across vagueness. Many problems can be very well described with the help of stochastic methods but may not cover fuzziness. For example, the point in time when an amendment will be implemented can be fuzzy. In terms of claims reserving the "fuzzification" of the volume in the additive method could be thinkable. The nature of this additional information is externally given information like e.g. from experts or information from comparable portfolios, market statistics, strategic business plan, premiums or number of contracts (cf. Wüthrich and Merz 2008, p. 23).

Up to now, reserving methods take into account either stochastic randomness or vagueness/fuzziness. Therefore, a next step should be the consideration of the combination of both sources of uncertainty. This problem might be solved with the help of fuzzy random variables. That concept may also allow for modeling dependencies between calendar years which have been left out when applying FNs. So far, fuzzy random variables only found their way into life insurance. Because of their ability to consider randomness and fuzziness they can be attractive for non-life insurance mathematics as well.

According to the requirements of Solvency II and the Swiss Solvency Test (SST) the one-year reserve risk needs to be quantified. The one-year reserve risk has not been reflected so far in the fuzzy literature. Besides that, the consideration of inflation effects (except for a first approach by Andrés Sánchez (2014), cf. Section 4.6.4) has been left out.

Current reserving methods like the PIC model (cf. Merz and Wüthrich 2010) incorporate two sources of data, i.e. paid and incurred data. As both data is often available it is reasonable to use all accessible useful information. Developing a procedure which allows for the incorporation of both paid and incurred data as well as fuzzy methods might be of interest.

A | Statistical Basics

This chapter deals with the representation of the statistical basics and notations used throughout the work, especially the multiple linear regression model and the important results are provided.

The classical multiple linear regression model is defined in the following way (cf. Fahrmeir et al. 2007, p. 62).

Definition A.1 (Classical multiple linear regression model)
Let $\mathbf{y} := (y_1, \ldots, y_n)^T \in \mathbb{R}^n$, $\mathbf{x}_i = (1, x_{i,1}, \ldots, x_{i,k})^T \in \mathbb{R}^{k+1}$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)^T \in \mathbb{R}^{k+1}$, $\mathbf{X} = (\mathbf{x}_1^T, \ldots, \mathbf{x}_n^T)^T \in \mathbb{R}^{n\times(k+1)}$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^T \in L^2(\mathbb{R}^n)$, $n \geq k+1$ *and* $i = 1, \ldots, n$. *That is,* $\boldsymbol{\varepsilon}$ *is a square-integrable random variable. The matrix* $\mathbf{X}$ *is called design matrix and* $\boldsymbol{\beta}$ *is referred to as vector of regression coefficients.*
The linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

is called classical multiple linear regression model if the following assumptions are fulfilled:

(i) $E[\boldsymbol{\varepsilon}] = \mathbf{0}$

(ii) $\mathrm{Cov}(\boldsymbol{\varepsilon}) = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T] = \sigma^2\mathbf{I}$, *where* $\sigma^2 > 0$

(iii) $\mathrm{rg}(\mathbf{X}) = k+1$, *i.e.* $\mathbf{X}$ *has full rank*

Under the additional assumption that the error term $\boldsymbol{\varepsilon}$ *is normally distributed, i.e.* $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, *we speak of a classical normal regression model.*

To estimate the regression coefficients $\boldsymbol{\beta}$ and the variance parameter σ^2 the method of ordinary least squares (OLS) is applied. In this method, the sum of squared residuals is minimized, i.e. the following function is considered (cf. Fahrmeir et al. 2007, p. 90)

$$\mathrm{OLS}(\boldsymbol{\beta}) := (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \sum_{i=1}^{n}(y_i - \mathbf{x}_i^T\boldsymbol{\beta})^2 = \sum_{i=1}^{n}\varepsilon_i^2 = \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}.$$

Then, the OLS-estimator $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \dots, \hat{\beta}_k)$ is defined as

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^{k+1}}{\arg\min}\,\mathrm{OLS}(\boldsymbol{\beta}).$$

Solving the minimization problem yields the representation (cf. Fahrmeir et al. 2007, p. 92)

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}.$$

The OLS-estimator $\hat{\boldsymbol{\beta}}$ is unbiased (cf. Fahrmeir et al. 2007, p. 101) and the Gauß-Markov-Theorem yields that it has least variances among all affine linear estimators (cf. Fahrmeir et al. 2007, p. 103). Thus, we speak of $\hat{\boldsymbol{\beta}}$ as a BLU (**B**est **L**inear **U**nbiased)-estimator.

For the construction of a confidence interval for a single regression coefficient β_j under the assumption of a normal distribution the t-statistics $t = \frac{\hat{\beta}_j - \beta_j}{\mathrm{se}(\hat{\beta}_j)}$ is used which follow a t-distribution with $n-(k+1)$ degrees of freedom. Then, a $(1-\alpha)$-confidence interval, $\alpha \in (0,1)$, is given by

$$[\hat{\beta}_j - t_{n-(k+1),1-\frac{\alpha}{2}}\mathrm{se}\left(\hat{\beta}_j\right), \hat{\beta}_j + t_{n-(k+1),1-\frac{\alpha}{2}}\mathrm{se}\left(\hat{\beta}_j\right)],$$

where $t_{n-(k+1),1-\frac{\alpha}{2}}$ is the $(1-\frac{\alpha}{2})$-quantile for a t-distributed random variable with $n-(k+1)$ degrees of freedom (cf. Fahrmeir et al. 2007, pp. 119 ff.).

In the classical multiple linear model uncorrelated and homoscedastic error terms $\boldsymbol{\varepsilon}$ are assumed. In case correlated and heteroscedastic errors are also allowed, one is in the context of the general multiple linear regression model (cf. Fahrmeir et al. 2007, p. 125).

Definition A.2

Let $\mathbf{W} \in \mathbb{R}^{n \times n}$ *be positive definite,* $\mathbf{y} := (y_1, \ldots, y_n)^T \in \mathbb{R}^n$, $\mathbf{x}_i = (1, x_{i,1}, \ldots, x_{i,k})^T \in \mathbb{R}^{k+1}$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)^T \in \mathbb{R}^{k+1}$, $\mathbf{X} = (\mathbf{x}_1^T, \ldots, \mathbf{x}_n^T)^T \in \mathbb{R}^{n \times (k+1)}$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^T \in L^2(\mathbb{R}^n)$, $n \geq k+1$ *and* $i = 1, \ldots, n$. *The linear model*

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

is called general multiple linear model if the following assumptions are satisfied:

(i) $E[\boldsymbol{\varepsilon}] = \mathbf{0}$

(ii) $\mathrm{Cov}(\boldsymbol{\varepsilon}) = E[\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}] = \sigma^2 \mathbf{W}$

(iii) $\mathrm{rg}(\mathbf{X}) = k+1$

Under the additional assumption that the error term $\boldsymbol{\varepsilon}$ *is normally distributed, i.e.* $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$, *we speak of a general normal regression model.*

The linear model as specified in Definition A.2 is also called weighted linear regression (with "weighting" matrix $\mathbf{W}$). Let $\mathbf{W}^{-\frac{1}{2}}$ denote the root of the matrix $\mathbf{W}^{-1}$, i.e. $\mathbf{W}^{-\frac{1}{2}}\mathbf{W}^{-\frac{1}{2}} = \mathbf{W}^{-1}$. By defining

$$\mathbf{y}^* := \mathbf{W}^{-\frac{1}{2}}\mathbf{y}$$
$$\mathbf{X}^* := \mathbf{W}^{-\frac{1}{2}}\mathbf{X}$$
$$\boldsymbol{\varepsilon}^* := \mathbf{W}^{-\frac{1}{2}}\boldsymbol{\varepsilon}$$

a general multiple linear model can be transformed to a classical multiple linear model

$$\mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}^*$$

(cf. Fahrmeir et al. 2007, pp. 125 f.). Hence, an estimator for the regression coefficient $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{W}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{-1} \mathbf{y} \tag{A.1}$$

and called Aitken-estimator (cf. Fahrmeir et al. 2007, p. 126).

Zusammenfassung

Die vorliegende Dissertation beschäftigt sich mit der Anwendung der Theorie unscharfer Mengen (fuzzy set theory) in der Schadenreservierung, einem Teilgebiet der Aktuarwissenschaften.

Aktuare innerhalb der Schadenreservierung befassen sich u.a. mit den folgenden zwei Aufgabenstellungen: der Prognose der Rückstellungen für zukünftige ausstehende Schadenzahlungen sowie der Quantifizierung von deren Risiko. Die Schadenreserve bildet einen großen Posten auf der Passivseite der Bilanz eines Versicherungsunternehmens, so dass aus Solvenz- und Ergebnisgründen weder eine Unter- noch eine Überschätzung empfehlenswert ist. Innerhalb der Schadenreservierung existieren eine Vielzahl von Verfahren: rein rechnerische, stochastische, verteilungsfreie, verteilungsgebundene, etc. Hierunter sind das Chain-Ladder- sowie das Bornhuetter Ferguson-Verfahren als populäre Verfahren zu nennen.

Unberücksichtigt bei den Verfahren bleibt häufig die Betrachtung von möglicher Vagheit und subjektiven Einschätzungen. Vagheit ist auch im Alltag, wie beispielsweise in der Sprache, allgegenwärtig. Die Theorie der unscharfen Mengen, welche im Jahre 1965 von Lotfi A. Zadeh begründet wurde, bietet einen formalen Ansatz subjektive Einschätzungen abzubilden. Erste Anwendungen in der Versicherungswissenschaft sind unseres Wissens nach seit den 1980er Jahren zu finden.

In der Dissertation werden zwei Ziele verfolgt: Zum Einen wird ein Überblick über die Anwendung von unscharfen Methoden in der Schadenreservierung dargestellt. Darüber hinaus werden in drei Kapiteln Wege aufgezeigt, wie subjektives Ermessen in Anwendungsmethoden in der Schadenreservierung in das Chain-Ladder- sowie das Bornhuetter Ferguson-Verfahren implementiert werden können.

Dem ersten Ansatz, dem Fuzzy Chain-Ladder-Verfahren, liegt die Beobachtung zu Grunde, dass es Fälle in der Praxis geben kann, in der Aktuare möglicherweise dazu tendieren, zuvor berechnete Abwicklungsfaktoren auf Grund ihrer subjektiven Einschätzungen anzupassen. Hierzu werden unscharfe Zahlen und deren Arithmetik herangezogen. Im zweiten Ansatz wird die Eigenschaft des Chain-Ladder Modells genutzt, dass es als Folge von linearen Modellen dargestellt werden kann. Hier wird eine Darstellung mit Hilfe von fuzzy Regressionsmodellen aufgezeigt, die Aktuaren die

Möglichkeit gewährt, den Informationsgehalt der Daten einzuschätzen und subjektiv zu bewerten. Der dritte Ansatz bezieht sich auf das Bornhuetter Ferguson-Verfahren. Diese Methode verwendet zusätzlich a priori-Informationen, welche entweder einer externen Datenquelle entstammen oder intern gegeben sind. Je nachdem welche Datenquelle den a priori Daten zu Grunde liegt, können diese mit Vagheit behaftet sein. Das vorgestellte Fuzzy Bornhuetter Ferguson-Verfahren, bietet einen Ansatz diese Vagheit mit Hilfe von unscharfen Zahlen abzubilden.

Allen Ansätzen ist gemein, dass sie nicht nur die ausstehenden Schadenzahlungen prognostizieren, sondern ebenso einen Ansatz zur Defuzzifikation sowie zur Quantifizierung der Unsicherheit der Prognose bieten.

Summary

This dissertation covers applications of fuzzy set theory in claims reserving, a field in actuarial science.

Actuaries working in claims reserving are faced, among others, with the following two tasks: the prediction of future outstanding loss liabilities, as well as the quantification of their risk. The claims reserve is a huge figure on the liabilities side of an insurer's balance sheet. Therefore, neither underestimating nor overestimating is advisable due to its impact on either solvency or profit. Within claims reserving there exist various methods whether they are only computational, stochastic, distribution-free, distributional, etc. Among those the chain-ladder and the Bornhuetter Ferguson method are probably the most common ones.

Often vagueness and subjective judgments have not been considered within these approaches. However, vagueness is omnipresent in our daily lives, e.g. in spoken language. The methodology of fuzzy sets has been introduced by Lotfi A. Zadeh in his seminal paper in 1965. It offers a formal approach to implement subjective assessments. To our best knowledge first applications of fuzzy set theory in insurance have been published in the 1980's.

This dissertation pursues two goals: First, an overview of applications of fuzzy set theory in claims reserving is presented. Moreover, three chapters present ways of how subjective assessment can be implemented in the chain-ladder as well as the Bornhuetter Ferguson method.

The first approach is the fuzzy chain-ladder method. It is based on the observation that there might be situations in practice in which actuaries tend to adjust previously calculated development factors due to their subjective judgment. For this method fuzzy numbers and their corresponding arithmetic are used. The second approach makes use of the representation of the chain-ladder model as a sequence of linear models. Here, a presentation using fuzzy regression models is shown. The advantage of this method for actuaries is the possibility to assess the information contained in the data as well as the feasibility for subjective adjustments. The third approach refers to the Bornhuetter Ferguson method. This technique additionally uses a priori information which can originate from an external or internal source. Depending on the source of the data they

can be afflicted with vagueness. The presented fuzzy Bornhuetter Ferguson method offers an approach to implement vagueness with the help of fuzzy numbers.

All approaches have in common that they not only yield predictions for the claims reserve, but also attempt to defuzzify the reserve and to quantify the predictions' uncertainty.

Bibliography

Abdullah, Lazim and Mohd Nordin Abd Rahman (2012). “Employee Likelihood of Purchasing Health Insurance Using Fuzzy Inference System”. In: *International Journal of Computer Science Issues* 9.1, pp. 112–116.

Alai, Daniel H., Michael Merz, and Mario V. Wüthrich (2009). “Mean Square Error of Prediction in the Bornhuetter–Ferguson Claims Reserving Method”. In: *Annals of Actuarial Science* 4.01, pp. 7–31.

Alai, Daniel H., Michael Merz, and Mario V. Wüthrich (2011). “Prediction uncertainty in the Bornhuetter-Ferguson claims reserving method: revisited”. In: *Annals of Actuarial Science* 5.01, pp. 7–17.

Altrock, Constantin von (1995). *Fuzzy Logic*. German. Second edition. Vol. 1: Technologie. R Oldenbourg.

Andrés Sánchez, Jorge de (2006). “Calculating Insurance Claim Reserves with Fuzzy Regression”. In: *Fuzzy Sets and Systems* 157.23, pp. 3091–3108.

Andrés Sánchez, Jorge de (2007). “Claim Reserving with Fuzzy Regression and Taylors’s Geometric Seperation Method”. In: *Insurance: Mathematics and Economics* 40.1, pp. 145–163.

Andrés Sánchez, Jorge de (2012). “Claim Reserving with Fuzzy Regression and the Two Ways of ANOVA”. In: *Applied Soft Computing* 12.8, pp. 2435–2441.

Andrés Sánchez, Jorge de (2014). “Fuzzy Claim Reserving in Non-Life Insurance”. In: *Computer Science and Information Sciences* 11.2, pp. 825–838.

Andrés Sánchez, Jorge de and Antonio Terceño Gómez (2003). “Applications of Fuzzy Regression in Actuarial Science”. In: *The Journal of Risk and Insurance* 70.4, pp. 665–699.

Bahrami, Tahereh and Masuod Bahrami (2015). “Claim Reserving with Fuzzy Regression”. In: *Cumhuriyet Science Journal* 36.3, pp. 704–709.

Bandemer, Hans and Siegfried Gottwald (1993). *Einführung in Fuzzy-Methoden.* German. Fourth edition. Akademie Verlag.

Bárdossy, András (1990). “Note on Fuzzy Regression”. In: *Fuzzy Sets and Systems* 37.1, pp. 65–75.

Barnett, Glen and Ben Zehnwirth (2000). “Best Estimates for Reserves”. In: *Proceedings of the Casualty Actuarial Society.* Vol. 87, pp. 245–321.

Başer, Furkan and Ayşen Apaydin (2010). “Hybrid Fuzzy Least-Squares Regression Analysis in Claims Reserving with Geometric Separation Method”. In: *Insurance: Mathematics and Economics* 47.2, pp. 113–122.

Başer, Furkan et al. (2011). “An Approach of Adaptive Network Based Fuzzy Inference System to Risk Classification in Life Insurance”. In: *Operations Research Proceedings 2010: Selected Papers of the Annual International Conference of the German Operations Research Society.* Ed. by Bo Hu et al. Springer, pp. 27–32.

Bellman, Richard and Magnus Giertz (1973). “On the Analytic Formalism of the Theory of Fuzzy Sets”. In: *Information Science* 5, pp. 149–156.

Benjamin, S. and L. M. Eagles (1986). “Reserves in Lloyd’s and the London Market”. In: *Journal of the Institute of Actuaries* 113.2, pp. 197–256.

Berry-Stölzle, Thomas R., Marie-Claire Koissi, and Arnold F. Shapiro (2010). “Detecting Fuzzy Relationships in Regression Models: The Case of Insurer Solvency Surveillance in Germany”. In: *Insurance: Mathematics and Economics* 46.3, pp. 554–567.

Bezdek, James C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms.* Kluwer Academic Publishers.

Black, Max (1937). “Vagueness. An Exercise in Logical Analysis”. In: *Philosophy of Science* 4.4, pp. 427–455.

Bonissone, Piero P. (2003). “The Life Cycle of a Fuzzy Knowledge-Based Classifier”. In: *22nd International Conference of the North American Fuzzy Information Processing Society, NAFIPS 2003.* IEEE, pp. 488–494.

Bonissone, Piero P. (2004). “Development and Maintenance of Fuzzy Models in Financial Applications”. In: *Soft Methodology and Random Information Systems.* Springer, pp. 50–66.

Bonissone, Piero P., Anil Varma, and Kareem S. Aggour (2005). “An Evolutionary Process for Designing and Maintaining a Fuzzy Instance-based Model (FIM)”. In: *Proceedings of the First Workshop of Genetic Fuzzy Systems (GFS 2005), Granada, Spain.*

Bornhuetter, Ronald L. and Ronald E. Ferguson (1972). “The Actuary and IBNR”. In: *Proceedings of the Casualty Actuarial Society* LIX, pp. 181–195.

Bosch, Harald (1993). “Entscheidung und Unschärfe: Eine entscheidungstheoretische Analyse der Fuzzy-Set-Theorie”. German. PhD thesis. Universität Hohenheim.

Brosius, Eric (1993). *Loss Development Using Credibility*. Casualty Actuarial Society Study Note. URL: `http://www.casact.org/library/studynotes/brosius6.pdf` (visited on 04/16/2016).

Buchwalder, Markus et al. (2006). “The Mean Square Error of Prediction in the Chain Ladder Reserving Method”. In: *ASTIN Bulletin* 36.2, pp. 521–542.

Buckley, James J. (2006). *Fuzzy Probability and Statistics*. Studies in Fuzziness and Soft Computing. Springer.

Buckley, James J. and Yunxia Qu (1990). “On Using α-Cuts to Evaluate Fuzzy Equations”. In: *Fuzzy Sets and Systems* 38.3, pp. 309–312.

Campos Ibáñez, Luis Miguel de and Antonio González Muñoz (1989). “A Subjective Approach for Ranking Fuzzy Numbers”. In: *Fuzzy Sets and Systems* 29.2, pp. 145–153.

Cantor, Georg (1895). “Beiträge zur Begründung der transfiniten Mengenlehre”. German. In: *Mathematische Annalen* 46.4, pp. 481–512.

Carlsson, Christer (1984). “On the Relevance of Fuzzy Sets in Management Science Methodology”. In: *Fuzzy Sets and Decision Analysis*. Ed. by Hans-Jürgen Zimmermann, Lotfi A. Zadeh, and Brian R. Gaines. North-Holland, pp. 11–28.

Carretero, Raquel Caro and Angel Sarabia Viejo (2000). “A Bonus-Malus System in the Fuzzy Set Theory [insurance pricing decisions]”. In: *The Ninth IEEE International Conference on Fuzzy Systems FUZZ IEEE 2000*. Vol. 2, pp. 1033–1036.

Casanovas, Montserrat, Agustín Torres Martínez, and José M. Merigó (2015). “Decision Making Processes of Non-Life Insurance Pricing Using Fuzzy Logic and OWA Operators”. In: *Economic Computation and Economic Cybernetics Studies and Research / Academy of Economic Studies* 49.2, pp. 1–19.

Celmiņš, Aivars (1987a). “Least Squares Model Fitting to Fuzzy Vector Data”. In: *Fuzzy Sets and Systems* 22.3, pp. 245–269.

Celmiņš, Aivars (1987b). “Multidimensional Least-Squares Fitting of Fuzzy Models”. In: *Mathematical Modelling* 9.9, pp. 669–690.

Chang, Yun-Hsi O. (2001). “Hybrid Fuzzy Least-Squares Regression Analysis and its Reliability Measures”. In: *Fuzzy Sets and Systems* 119.2, pp. 225–246.

Chang, Yun-Hsi O. and Bilal M. Ayyub (2001). "Fuzzy Regression Methods – A Comparative Assessment". In: *Fuzzy Sets and Systems* 119.2, pp. 187–203.

Collan, Mikael, Mario Fedrizzi, and Pasi Luukka (2016). "Possibilistic Risk Aversion in Group Decisions: Theory with Application in the Insurance of Giga-Investments Valued through the Fuzzy Pay-Off Method". In: *Soft Computing*, pp. 1–12.

Culp, Christopher L. (2006). *Structured Finance and Insurance: the ART of Managing Capital and Risk*. Wiley.

Cummins, J. David and Richard A. Derrig (1997). "Fuzzy Financial Pricing of Property-Liability Insurance". In: *North American Actuarial Journal* 1.4, pp. 21–40.

Dahms, René (2008). "A Loss Reserving Method for Incomplete Claim Data". In: *Bulletin Swiss Association of Actuaries*, pp. 127–148.

Dahms, René (2012). "Linear Stochastic Reserving Methods". In: *ASTIN Bulletin* 42.1, pp. 1–34.

De Luca, Aldo and Settimo Termini (1972). "A Definition of a Nonprobabilistic Entropy in the Setting of Fuzzy Sets Theory". In: *Information and Control* 20.4, pp. 301–312.

Derrig, Richard A. and Krzysztof M. Ostaszewski (1995). "Fuzzy Techniques of Pattern Recognition in Risk and Claim Classification". In: *The Journal of Risk and Insurance* 62.3, pp. 447–482.

Derrig, Richard A. and Krzysztof M. Ostaszewski (1999). "Fuzzy Sets Methodologies in Actuarial Science". In: *Practical Applications of Fuzzy Technologies*. Ed. by Hans-Jürgen Zimmermann. Vol. 6. The Handbooks of Fuzzy Sets Series. Springer, pp. 531–553.

Diamond, Phil (1988). "Fuzzy Least Squares". In: *Information Sciences* 46.3, pp. 141–157.

Diamond, Phil (1992). "Least Squares and Maximum Likelihood Regression for Fuzzy Linear Models". In: *Fuzzy Regression Analysis*. Ed. by Janusz Kacprzyk and Mario Fedrizzi. Physica, pp. 137–151.

Diamond, Phil and Peter E. Kloeden (1994). *Metric Spaces of Fuzzy Sets: Theory and Applications*. World Scientific.

Diamond, Phil and Hideo Tanaka (1998). "Fuzzy Regression Analysis". In: *Fuzzy Sets in Decision Analysis, Operations Research and Statistics*. Ed. by Roman Słowiński. Kluwer Academic Publishers.

Dubois, Didier (1983). "Modèles Mathematiques de L'Imprecis et de L'Incertain en Vue de L'Applications aux Techniques d'Aide à la Decision". French. PhD thesis. Université Joseph-Fourier – Grenoble I; Institut National Polytechnique de Grenoble.

Dubois, Didier, Hung T. Nguyen, and Henri Prade (2000). "Possibility Theory, Probability and Fuzzy Sets: Misunderstandings, Bridges and Gaps". In: *Fundamentals of Fuzzy Sets*. Ed. by Didier Dubois and Henri Prade. Vol. 7. The Handbooks of Fuzzy Sets Series. Springer, pp. 343–438.

Dubois, Didier and Henri Prade (1978). "Operations on Fuzzy Numbers". In: *International Journal of Systems Science* 9.6, pp. 613–626.

Dubois, Didier and Henri Prade (1979). "Fuzzy Real Algebra: Some Results". In: *Fuzzy Sets and Systems* 2.4, pp. 327–348.

Dubois, Didier and Henri Prade (1980). *Fuzzy Sets and Systems: Theory and Applications*. Academic Press.

Dubois, Didier and Henri Prade (1993). "Fuzzy Numbers: An Overview". In: *Readings in Fuzzy Sets for Intelligent Systems*. Ed. by Didier Dubois, Henri Prade, and Ronald Yager. Morgan Kaufmann.

Dubois, Didier and Henri Prade (1994). "Fuzzy Sets – A Convenient Fiction for Modeling Vagueness and Possibility". In: *IEEE Transactions on Fuzzy Systems* 2.1, pp. 16–21.

Dubois, Didier and Henri Prade (2000). *Fundamentals of Fuzzy Sets*. Vol. 7. The Handbooks of Fuzzy Sets Series. Springer.

Ebanks, Bruce R. (1983). "On Measures of Fuzziness and their Representations". In: *Journal of Mathematical Analysis and Applications* 94, pp. 24–37.

Ebanks, Bruce, Waldemar Karwowski, and Krzysztof Ostaszewski (1992). "Application of Measures of Fuzziness to Risk Classification in Insurance". In: *Fourth International Conference on Computing and Information, 1992. Proceedings*, pp. 290–291.

Eberhart, Russell C. and Yuhui Shi (2011). *Computational Intelligence: Concepts to Implementations*. Elsevier.

England, Peter D. and Richard J. Verrall (2002). "Stochastic Claims Reserving in General Insurance". In: *British Actuarial Journal* 8.3, pp. 443–518.

Erbach, David W. and Eric Seah (1993). "Discussion of the Application of Fuzzy Sets to Group Health Underwriting by Virginia R. Young". In: *Transactions of the Society of Actuaries* 45, pp. 585–587.

Fahrmeir, Ludwig, Thomas Kneib, and Stefan Lang (2007). *Regression*. German. Statistik und ihre Anwendungen. Springer.

Finetti, Bruno de (1981). *Wahrscheinlichkeitstheorie. Einführende Synthese mit kritischem Anhang*. German. R. Oldenbourg.

Forschner, Markus (1998). *Prozessorientiertes Investitionscontrolling: Bewertung von Informationssystemen mit Hilfe der Fuzzy Logic*. German. Deutscher Universitäts Verlag.

Gençtürk, Yasemin et al. (2011). "Fuzzy Pricing of High Excess Loss Layer When Modeling the Tail with Generalized Pareto Distribution". In: *İstatistikçiler Dergisi: İstatistik ve Aktüerya* 4.1, pp. 16–30.

Gil, María Angeles, Miguel López-Díaz, and Dan A. Ralescu (2006). "Overview on the Development of Fuzzy Random Variables". In: *Fuzzy Sets and Systems* 157.19, pp. 2546–2557.

Gisler, Alois and Mario V. Wüthrich (2008). "Credibility for the Chain Ladder Reserving Method". In: *ASTIN Bulletin* 38.2, pp. 565–600.

Goovaerts, Marc et al. (1990). *Effective Actuarial Methods*. North-Holland.

Hanss, Michael (2005). *Applied Fuzzy Arithmetic – An Introduction with Engineering Applications*. Springer.

Harnek, R. F. (1966). "Formula Loss Reserves". In: *Insurance Accounting and Statistical Proceedings*.

Heberle, Jochen and Anne Thomas (2014). "Combining Chain-Ladder Claims Reserving with Fuzzy Numbers". In: *Insurance: Mathematics and Economics* 55, pp. 96–104.

Heberle, Jochen and Anne Thomas (2016). "The Fuzzy Bornhuetter-Ferguson Method – An Approach with Fuzzy Numbers". In: *Annals of Actuarial Science*, forthcoming.

Hess, Klaus T. and Klaus D. Schmidt (2002). "A Comparison of Models for the Chain-Ladder Method". In: *Insurance: Mathematics and Economics* 31.3, pp. 351–364.

Holtfort, Thomas (2011). "Intuition als Managementfaktor im Finanzdienstleistungssektor". German. In: *Versicherungswirtschaft* 7, pp. 507–511.

Horgby, Per-Johan (1998). "Risk Classification by Fuzzy Inference". In: *The Geneva Papers on Risk and Insurance Theory* 23.1, pp. 63–82.

Horgby, Per-Johan, Ralf Lohse, and Nicola-Alexander Sittaro (1997). "Fuzzy Underwriting: An Application of Fuzzy Logic to Medical Underwriting". In: *Journal of Actuarial Practice* 5.1, p. 79.

Imriyas, Kamardeen (2009). "An Expert System for Strategic Control of Accidents and Insurers' Risks in Building Construction Projects". In: *Expert Systems with Applications* 36.2, pp. 4021–4034.

Imriyas, Kamardeen, Low Sui Pheng, and Evelyn Ai-Lin Teo (2007). "A Fuzzy Knowledge-Based System for Premium Rating of Workers' Compensation Insurance for Building Projects". In: *Construction Management and Economics* 25.11, pp. 1177–1195.

Imriyas, Kamardeen et al. (2006). "A Fuzzy Expert System for Computing Workers' Compensation Insurance Premiums in Construction: A Conceptual Framework". In: *Architectural Science Review* 49.3, pp. 270–284.

Institute of Actuaries (1997). *Claims Reserving Manual*. URL: `https://www.actuaries.org.uk/documents/claims-reserving-manual-volume-1` (visited on 04/16/2016).

Ishibuchi, Hisao and Manabu Nii (2001). "Fuzzy Regression Using Asymmetric Fuzzy Coefficients and Fuzzified Neural Networks". In: *Fuzzy Sets and Systems* 119.2, pp. 273–290.

Kaas, Rob et al. (2008). *Modern Actuarial Risk Theory: Using R*. Second Edition. Springer.

Kacprzyk, Janusz and Mario Fedrizzi, eds. (1992). *Fuzzy Regression Analysis*. Physica.

Kaufmann, Arnold (1975). *Introduction to the Theory of Fuzzy Subsets. Vol. I: Fundamental Theoretical Elements*. Academic Press.

Kerkez, Marija (2013a). "Estimating Reserve for Incurred but not Reported (IBNR) Claims with Fuzzy Regression Model". In: *Advances in Accounting, Auditing and Risk Management*. Ed. by Vladimir Marascu-Klein. Proceedings of the 2nd International Conference on Finance, Accounting and Auditing (FAA '13), Brasov, Romania, pp. 166–170. URL: `http://www.wseas.us/e-library/conferences/2013/Brasov/FARI/FARI-28.pdf` (visited on 04/16/2016).

Kerkez, Marija (2013b). "Fuzzy Regression Model Applied to the Incurred but not Reported Claims (IBNR)". In: *Recent Researchers in Law Science and Finances*. Ed. by Kalliopi Kalampouka and Carmen Nastase. Proceedings of the 4th International Conference on Finance, Accounting and Law (ICFA '13), Chania, Greece, pp. 216–221. URL: `http://www.wseas.us/e-library/conferences/2013/Chania/ICFA/ICFA-32.pdf` (visited on 04/16/2016).

Kim, Joseph H. T. and Joocheol Kim (2014). "Fuzzy Regression Towards a General Insurance Application". In: *Journal of Applied Mathematics & Informatics* 32.3-4, pp. 343–357.

Kolmogorov, Andrey Nikolaevich (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. German. Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer.

Körner, Ralf and Wolfgang Näther (1998). "Linear Regression with Random Fuzzy Variables: Extended Classical Estimates, Best Linear Estimates, Least Squares Estimates". In: *Journal of Information Sciences* 109, pp. 95–118.

Kosko, Bart (1990). "Fuzziness vs. Probability". In: *International Journal of General Systems* 17.2-3, pp. 211–240.

Krätschmer, Volker (2001). "A Unified Approach to Fuzzy Random Variables". In: *Fuzzy Sets and Systems* 123.1, pp. 1–9.

Kremer, Erhard (1982). "IBNR-Claims and the Two-Way Model of ANOVA". In: *Scandinavian Actuarial Journal* 1982.1, pp. 47–55.

Kruse, Rudolf, Jörg Gebhardt, and Frank Klawonn (1995). *Fuzzy-Systeme*. German. Second edition. B. G. Teubner.

Kruse, Rudolf and Klaus Dieter Meyer (1987). *Statistics with Vague Data*. Reidel.

Kwakernaak, Huibert (1978). "Fuzzy Random Variables—I. Definitions and Theorems". In: *Information Sciences* 15.1, pp. 1–29.

Kwakernaak, Huibert (1979). "Fuzzy Random Variables—II. Algorithms and Examples for the Discrete Case". In: *Information Sciences* 17.3, pp. 253–278.

Lai, Li-Hua (2006). "Underwriting Profit Margin of P/L Insurance in the Fuzzy-ICAPM". In: *The GENEVA Risk and Insurance Review* 31.1, pp. 23–34.

Leekwijk, Werner van and Etienne E. Kerre (1999). "Defuzzification: Criteria and Classification". In: *Fuzzy Sets and Systems* 108.2, pp. 159–178.

Lemaire, Jean (1988). "Claims Reserving in Non-Life Insurance (Book Review)." In: *Journal of Risk and Insurance* 55.2, pp. 396–398.

Lemaire, Jean (1990). "Fuzzy Insurance". In: *ASTIN Bulletin* 20.1, pp. 33–55.

Liu, Ying, Xiaozhong Li, and Yinli Liu (2015). "The Bounds of Premium and Optimality of Stop Loss Insurance Under Uncertain Random Environments". In: *Insurance: Mathematics and Economics* 64, pp. 273–278.

Ludwig, Alexander, Christiane Schmeißer, and Katrin Thänert (2009). "Linear Models in Loss Reserving". In: *Dresdner Schriften zur Versicherungsmathematik* 1.

Luukka, Pasi and Mikael Collan (2015). "New Fuzzy Insurance Pricing Method for Giga-Investment Project Insurance". In: *Insurance: Mathematics and Economics* 65, pp. 22–29.

Mack, Thomas (1993). "Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates". In: *ASTIN Bulletin* 23.2, pp. 213–225.

Mack, Thomas (1999). "The Standard Error of Chain-Ladder Reserve Estimates, Recursive Calculation and Inclusion of a Tail Factor". In: *ASTIN Bulletin* 29.2, pp. 361–366.

Mack, Thomas (2000). "Credible Claims Reserve: The Benktander Method". In: *ASTIN Bulletin* 30.2, pp. 333–348.

Mack, Thomas (2002). *Schadenversicherungsmathematik*. German. Schriftenreihe angewandte Versicherungsmathematik. VVW, Verlag Versicherungswirtschaft.

Mack, Thomas (2008). "The Prediction Error of Bornhuetter/Ferguson". In: *ASTIN Bulletin* 38.1, pp. 87–103.

Menger, Karl (1951). "Ensembles flous et fonctions aléatoires". French. In: *Comptes Rendus de l'Académie des Sciences de Paris* 232, pp. 2001–2003.

Merz, Michael and Mario V. Wüthrich (2007). "Prediction Error of the Expected Claims Development Result in the Chain Ladder Method". In: *Bulletin of Swiss Association of Actuaries* 1, pp. 117–137.

Merz, Michael and Mario V. Wüthrich (2008). "Prediction Error of the Multivariate Chain Ladder Reserving Method". In: *North American Actuarial Journal* 12.2, pp. 175–197.

Merz, Michael and Mario V. Wüthrich (2010). "Paid–Incurred Chain Claims Reserving Method". In: *Insurance: Mathematics and Economics* 46.3, pp. 568–579.

Merz, Michael and Mario V. Wüthrich (2013a). "Estimation of Tail Development Factors in the Paid-Incurred Chain Reserving Method". In: *Variance* 7.1, pp. 61–73.

Merz, Michael and Mario V. Wüthrich (2013b). *Mathematik für Wirtschaftswissenschaftler: Die Einführung mit vielen ökonomischen Beispielen*. German. Vahlen.

Morgenstern, Oskar (1965). *Über die Genauigkeit wirtschaftlicher Beobachtungen*. German. Second Edition. Physica.

Moskowitz, Herbert and Kwangjae Kim (1993). "On Assessing the H Value in Fuzzy Linear Regression". In: *Fuzzy Sets and Systems* 58.3, pp. 303–327.

Murphy, Daniel (1994). "Unbiased Loss Development Factors". In: *Proceedings of the Casualty Actuarial Society* 81, pp. 154–222.

Onoghojobi, B. and N. P. Olewuezi (2015). "Addressing the Issue of Chain Ladder Forecasting Using Fuzzy Regression Model". In: *Advances in Fuzzy Sets and Systems* 20.1, pp. 7–24.

Ostaszewski, Krzysztof (1993). *An Investigation into Possible Applications of Fuzzy Set Methods in Actuarial Science*. Society of Actuaries.

Outreville, J. Francois (1988). *Theory and Practice of Insurance*. Springer.

Pal, Nikhil R. and James C. Bezdek (1994). “Measuring Fuzzy Uncertainty”. In: *IEEE Transactions on Fuzzy Systems* 2.2, pp. 107–118.

Peters, Gareth W., Mario V. Wüthrich, and Pavel V. Shevchenko (2010). “Chain Ladder Method: Bayesian Bootstrap Versus Classical Bootstrap”. In: *Insurance: Mathematics and Economics* 47.1, pp. 36–51.

Peters, Georg (1994). “Fuzzy Linear Regression with Fuzzy Intervals”. In: *Fuzzy Sets and Systems* 63.1, pp. 45–55.

Petry, Nikolaus (1998). “Lineare Fuzzy-Regression”. German. Dissertation. Saarland University.

Pröhl, Carsten and Klaus D. Schmidt (2005). “Multivariate Chain-Ladder”. In: *Dresdner Schriften zur Versicherungsmathematik* 3.

Puri, Madan L. and Dan A. Ralescu (1986). “Fuzzy Random Variables”. In: *Journal of Mathematical Analysis and Applications* 114.2, pp. 409–422.

Quarg, Gerhard and Thomas Mack (2004). “Munich Chain Ladder”. In: *Blätter der DGVFM* 26.4, pp. 597–630.

Redden, David T. and William H. Woodall (1994). “Properties of Certain Fuzzy Linear Regression Methods”. In: *Fuzzy Sets and Systems* 64.3, pp. 361–375.

Runkler, Thomas A. and Manfred Glesner (1993). “A Set of Axioms for Defuzzification Strategies Towards a Theory of Rational Defuzzification Operators”. In: *Second IEEE International Conference on Fuzzy Systems 1993*. IEEE, pp. 1161–1166.

Russell, Bertrand (1923). “Vagueness”. In: *Australasian Journal of Psychology and Philosophy* 1.2, pp. 84–92.

Saluz, Annina, Alois Gisler, and Mario V. Wüthrich (2011). “Development Pattern and Prediction Error for the Stochastic Bornhuetter-Ferguson Claims Reserving Method”. In: *ASTIN Bulletin* 41.02, pp. 279–313.

Savic, Dragan A. and Witold Pedrycz (1991). “Evaluation of Fuzzy Linear Regression Models”. In: *Fuzzy Sets and Systems* 39.1, pp. 51–63.

Schmidt, Klaus D. and Michael Radtke (2004). *Handbuch zur Schadenreservierung*. German. Verlag Versicherungswirtschaft.

Schmidt-Gallas, Dirk (2014). “Preise sind Chefsache”. In: *Versicherungswirtschaft* 4, pp. 20–22.

Schneeweiss, Hans and Hans-Joachim Mittag (1986). *Lineare Modelle mit fehlerbehafteten Daten*. German. Physica.

Seising, Rudolf (2005). *Die Fuzzifizierung der Systeme: Die Entstehung der Fuzzy Set Theorie und ihrer ersten Anwendungen – ihre Entwicklung bis in die 70er Jahre des 20. Jahrhunderts*. German. Franz Steiner Verlag.

Shapiro, Arnold F. (2004). "Fuzzy Logic in Insurance". In: *Insurance: Mathematics and Economics* 35.2, pp. 399–424.

Shapiro, Arnold F. (2006). *Fuzzy Regression Models and the Term Structure of Interest Rates*. Actuarial Reasearch Clearing House 2006-1.

Shapiro, Arnold F. (2007). "An Overview of Insurance Uses of Fuzzy Logic". In: *Computational Intelligence in Economics and Finance*. Ed. by Shu-Heng Chen, Paul P. Wang, and Tzu-Wen Kuo. Springer, pp. 25–61.

Shapiro, Arnold F. (2009). "Fuzzy Random Variables". In: *Insurance: Mathematics and Economics* 44.2, pp. 307–314.

Shapiro, Arnold F. and Lakhmi C. Jain (2003). *Intelligent and Other Computational Techniques in Insurance: Theory and Applications*. Vol. 6. World Scientific.

Sherman, Richard E. (1984). "Extrapolating, Smoothing and Interpolating Development Factors". In: *Proceedings of the Casualty Actuarial Society* 71, pp. 122–155.

Straub, Erwin (1988). *Non-life Insurance Mathematics*. Springer.

Tanaka, Hideo (1987). "Fuzzy Data Analysis by Possibilistic Linear Models". In: *Fuzzy Sets and Systems* 24.3, pp. 363–375.

Tanaka, Hideo and Hisao Ishibuchi (1992). "A Possibilistic Regression Analysis Based on Linear Programming". In: *Fuzzy Regression Analysis*. Ed. by Janusz Kacprzyk and Mario Fedrizzi. Physica, pp. 47–60.

Tanaka, Hideo, Satoru Uejima, and Kiyoji Asai (1980). "Fuzzy Linear Regression Model". In: *Applied Systems and Cybernetics: Proceedings of the International Congress on Applied Systems Research and Cybernetics*. Ed. by George E. Lasker, pp. 12–16.

Tanaka, Hideo, Satoru Uejima, and Kiyoji Asai (1982). "Linear Regression Analysis with Fuzzy Model". In: *IEEE Transactions on Systems, Man and Cybernetics* 12.6, pp. 903–907.

Taylor, Greg C. (1977). "Separation of Inflation and other Effects from the Distribution of Non-Life Insurance Claim Delays". In: *ASTIN Bulletin* 9.1-2, pp. 219–230.

Taylor, Greg C. (1978). "Statistical Testing of a Non-Life Insurance Run-Off Model". In: *I.B.N.R. – Proceedings of the First Meeting of the Contact Group Actuarial Science*. Ed. by F. Etienne de Vylder and Marc Goovaerts, pp. 37–64.

Taylor, Greg C. (2000). *Loss Reserving: An Actuarial Perspective*. Kluwer Academic Publishers.

Taylor, Greg C. and Frank R. Ashe (1983). “Second Moments of Estimates of Outstanding Claims”. In: *Journal of Econometrics* 23.1, pp. 37–61.

Verrall, Richard J. (2004). “A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving”. In: *North American Actuarial Journal* 8.3, pp. 67–89.

Verrall, Richard J. and Yakoub H. Yakoubov (1999). “A Fuzzy Approach to Grouping by Policyholder Age in General Insurance”. In: *Journal of Actuarial Practice* 7, pp. 181–203.

Wagenknecht, Michael, Rainer Hampel, and Veit Schneider (2001). “Computational Aspects of Fuzzy Arithmetics Based on Archimedean t-Norms”. In: *Fuzzy Sets and Systems* 123.1, pp. 49–62.

Wang, Hsiao-Fan and Ruey-Chyn Tsaur (2000). “Insight of a Fuzzy Regression Model”. In: *Fuzzy Sets and Systems* 112.3, pp. 355–369.

Wit, G. Willem de (1982). “Underwriting and Uncertainty”. In: *Insurance: Mathematics and Economics* 1.4, pp. 277–285.

Wünsche, Andreas and Wolfgang Näther (2002). “Least-Squares Fuzzy Regression with Fuzzy Random Variables”. In: *Fuzzy Sets and Systems* 130.1, pp. 43–50.

Wüthrich, Mario V. and Michael Merz (2008). *Stochastic Claims Reserving Methods in Insurance*. Wiley.

Yakoubov, Yakoub H. and Steven Haberman (1998). *Review of Actuarial Applications of Fuzzy Set Theory*. Actuarial research paper 105. City University, Department of Actuarial Science and Statistics.

Yao, Kai and Zhongfeng Qin (2015). “A Modified Insurance Risk Process with Uncertainty”. In: *Insurance: Mathematics and Economics* 62, pp. 227–233.

Young, Virginia R. (1993). “The Application of Fuzzy Sets to Group Health Underwriting”. In: *Transactions of Society of Actuaries* 45, pp. 551–590.

Young, Virginia R. (1996). “Insurance Rate Changing: A Fuzzy Logic Approach”. In: *Journal of Risk and Insurance* 63.3, pp. 461–483.

Young, Virginia R. (1997). “Adjusting Indicated Insurance Rates: Fuzzy Rules that Consider both Experience and Auxiliary Data”. In: *Proceedings of the Casualty Actuarial Society*. Vol. 84, pp. 734–765.

Zadeh, Lotfi A. (1965). “Fuzzy Sets”. In: *Information and Control* 8.3, pp. 338–353.

Zadeh, Lotfi A. (1975a). "The Concept of a Linguistic Variable and its Application to Approximate Reasoning – Part I". In: *Information Sciences* 8.3, pp. 199–249.

Zadeh, Lotfi A. (1975b). "The Concept of a Linguistic Variable and its Application to Approximate Reasoning – Part II". In: *Information Sciences* 8.4, pp. 301–357.

Zadeh, Lotfi A. (1975c). "The Concept of a Linguistic Variable and its Application to Approximate Reasoning – Part III". In: *Information Sciences* 9.1, pp. 43–80.

Zadeh, Lotfi A. (1995). "Discussion: Probability Theory and Fuzzy Logic are Complementary Rather than Competitive". In: *Technometrics* 37.3, pp. 271–276.

Zadeh, Lotfi A. (1996). "Fuzzy Logic = Computing with Words". In: *IEEE Transactions on Fuzzy Systems* 4.2, pp. 103–111.

Zimmermann, Hans-Jürgen (1983). "Using Fuzzy Sets in Operational Research". In: *European Journal of Operational Research* 13.3, pp. 201–216.

Zimmermann, Hans-Jürgen (2001). *Fuzzy Set Theory and its Applications*. Fourth Edition. Kluwer Academic Publishers.

Zimmermann, Hans-Jürgen, Joachim Angstenberger, and Richard Weber (1999). "Intelligent Data Analysis and Fuzzy Control". In: *Fuzzy Algorithms for Control*. Ed. by Henk B. Verbruggen, Hans-Jürgen Zimmermann, and Robert Babuška. Springer, pp. 223–242.