

Prüfung klimadynamischer Mortalitätsmodelle mit externen Daten

Die Vorhersage individuen-basierter Baummortalitäten ist essenziell zur Abschätzung veränderter Risiken im Zuge des Klimawandels. Überlebenszeitmodelle bieten hierfür einen Modellierungsansatz. Ursprünglich entwickelt für medizinische Studien, finden Überlebenszeitmodelle ihre ersten Anwendungen nun auch in den Forstwissenschaften.

Um Aussagen bezüglich der Modellgenauigkeit und -verlässlichkeit zu erhalten, wird ein von der Bayerischen Landesanstalt für Wald und Forstwirtschaft (LWF) entwickeltes Überlebenszeitmodell mittels externer Daten evaluiert. Der vorliegende Beitrag zeigt hierzu Methoden und Ergebnisse auf.

Janet Maringer, Axel Albrecht

Natürliche Baummortalität ist ein wichtiger ökologischer Prozess, der durch den Vitalitätsverlust eines Baumes hervorgerufen wird. Dabei wird zwischen störungsbedingter (irregulärer) und konkurrenzbedingter (regulärer) Mortalität unterschieden [1]. Unabhängig von der Mortalitätsursache ist die Vorhersage individuenbasierter Baummortalitäten essenziell, um ein verändertes Risiko im Zuge des Klimawandels abschätzen zu können [2]. Allerdings ist es aufgrund der hohen Variabilität stochastisch auftretender Ereignisse und

deren im Verhältnis zur Lebensdauer von Bäumen geringen Anzahl eine statistische Herausforderung, Baummortalitäten zu modellieren [1].

Ein häufig gewählter statistischer Ansatz für die Modellierung von Baummortalitäten ist die logistische Regression [4, 5]. Dabei wird ein dichotomes (lebendig/tot) Ereignis über einen einheitlichen Zeitausschnitt betrachtet, z. B. die Mortalitätswahrscheinlichkeit für die nächsten fünf Jahre. Eine Alternative zu diesem Ansatz ist die Modellierung der Überlebenszeitwahrscheinlichkeit (ÜZW) unter Verwendung von Überlebenszeitmodellen mit zensierter Datenstruktur. Tritt das

Schneller Überblick

- Testen von Survival-Modellen mit externem Datensatz, basierend auf Langzeitdaten der FVA
- Geprüft wurden Modellstruktur und Effekte
- Von der LWF entwickelte Modelle liefern bei Anwendung auf externen Datensatz generell gute Modellperformanz
- Das Niveau von geschätzten Baumartenmortalitäten wird als realistisch geschätzt
- Implementierung von zeitveränderlichen Variablen in Modelle wird empfohlen



Abb. 1: Mortalität in Nadelholzbeständen wird häufig durch Sturm verursacht.

gewünschte Ereignis nicht bis zum Ende der Beobachtung ein, spricht man von zensierten Daten. Zusätzlich können die Daten trunziert sein. So wird von links trunzierten Daten gesprochen, wenn alle Objekte erst ab einem bestimmten Kriterium (bspw. alle Bäume ab dem Alter 60 Jahre) in die Studie mit aufgenommen werden.

Überlebenszeitmodelle lassen sich in drei große Gruppen (nicht-parametrische, semi-parametrische und parametrische Schätzer) unterteilen. Zur nicht-parametrischen Gruppe zählt der Kaplan-Meier-Schätzer, ein fundamentales Instrument der Lebenszeitanalyse. Dieser nicht-parametrische Schätzer bildet eine Basis für grafische Verfahren, mit deren Hilfe parametrische Modelle überprüft werden können. Zu den semi-parametrischen Schätzern gehört das Cox-Modell [6]. Hier bleibt das Basisrisiko über die Zeit konstant. Bei parametrischen Modellen – zu diesen zählen bspw. *Accelerated Failure Time Modelle* (AFT-Modelle) – wird der Überlebenszeit eine Verteilung zugrunde gelegt. Häufig sind es die Weibull- oder Lognormal-Verteilung. AFT-Modelle haben den Vorteil, dass der Effekt der Prädiktoren in absoluten Zahlen (bspw. ein Jahr) geschätzt wird [7].

Ursprünglich entwickelt für medizinische Studien, finden AFT-Modelle seit einigen Jahren Anwendung in den Forstwissenschaften [8, 9]. Der vorliegende Beitrag legt erstmals den Fokus auf die Evaluierung von AFT-Modellen und geht den Fragen nach, inwieweit die Modellstruktur und Effekte plausibel abgebildet werden.

Wie können wir AFT-Modelle testen?

Um Aussagen bezüglich der Modellgenauigkeit und -verlässlichkeit zu erhalten, ist eine Evaluierung prinzipiell

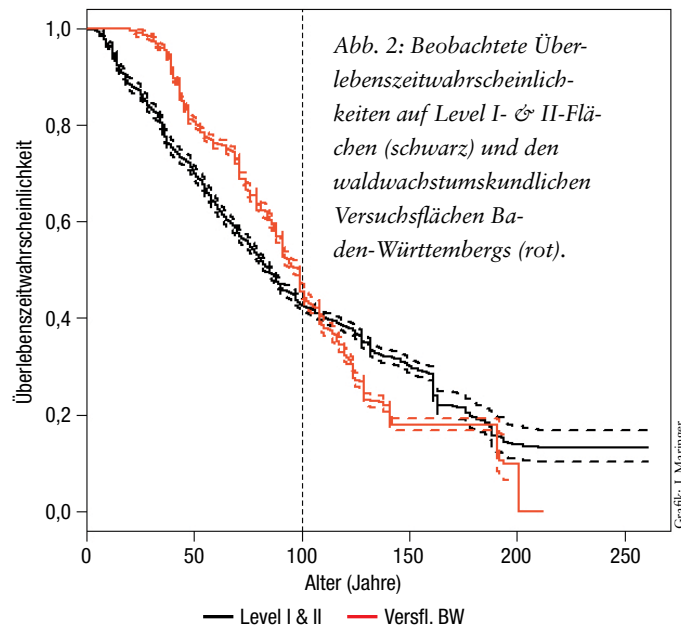
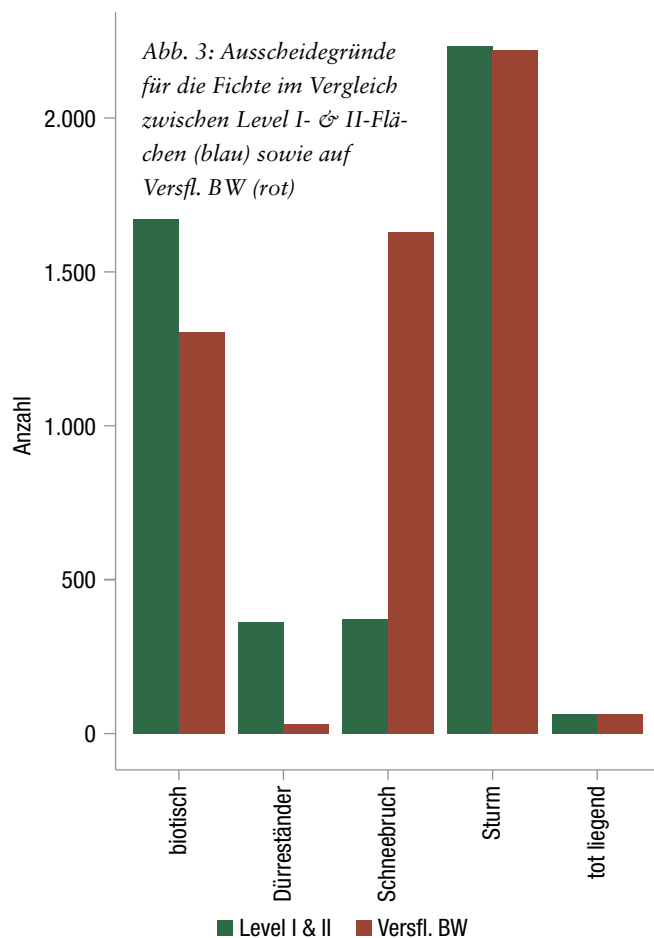


Abb. 3: Ausscheidegründe für die Fichte im Vergleich zwischen Level I- & II-Flächen (blau) sowie auf Versfl. BW (rot)



für jede Art von Modellen erforderlich [10]. Dies kann entweder durch interne, zeitliche oder externe Evaluierung geschehen. Bei der internen und zeitlichen Evaluierung wird ein Datensatz in einen Entwicklungs- und Evaluierungsdatensatz geteilt, während eine externe Evaluierung mittels eines neuen Datensatzes durchgeführt wird [11]. Im vorlie-

genden Beitrag wird ein von der LWF (Entwicklungsdatensatz basiert auf Europäischen Daten – Level I & II) entwickeltes AFT-Modell mit Linkstrunkierung mittels externer Daten evaluiert (s. den Beitrag ab S. 10 in dieser Ausgabe). Dabei wird auf einen räumlich und zeitlich sehr umfangreichen Datensatz der Forstlichen Versuchsanstalt Baden-Württemberg (FVA; Evaluierungsdatensatz: Versfl BW) zurückgegriffen. Der Evaluierungsdatensatz wurde auf den Entwicklungsdatensatz angepasst (Periode 1994 bis 2014) und umfasst Informationen zu 147.741 Bäumen mit einer Ausfallrate von 8,7 %.

Modellevaluierung Schritt für Schritt

Schritt I:
Datengrundlagen im Vergleich

Die von der LWF entwickelten Modelle wurden baumartenspezifisch kalibriert und evaluiert. Das Prinzip der Evaluierung von AFT-Modellen wird hier exemplarisch anhand der Fichte beschrieben. In einem ersten Schritt wurde dazu die grundlegende Datenstruktur des Entwicklungs- und Evaluierungsdatensatzes verglichen. Würden grundlegende Unterschiede zwischen den Datensätzen bestehen, könnte es zu Verzerrungen kommen oder andere Prädiktoren im Modell ausgewählt werden.

In einem ersten Schritt wurden die beobachteten ÜZW verglichen. Hierfür wurde der Kaplan-Meier-Schätzer ausgewählt. Dieser nicht-parametrische Schätzer berechnet die Wahrscheinlichkeit, dass innerhalb eines bestimmten Zeitintervalls ein Baum nicht ausfällt. Dazu wird das Verhältnis gebildet zwischen der Anzahl der lebenden Bäume und der Grundgesamtheit (sprich alle Bäume) zum Zeitpunkt t .

Dies wird mit dem zum Zeitpunkt $t-1$ berechneten Risiko multipliziert. Die berechnete Überlebenszeitkurve ist somit monoton sinkend.

Ein visueller Vergleich der ÜZW der Fichte auf Level I und II mit den Versfl. BW zeigt nur eine geringe Abweichung im zeitlichen Verlauf (Abb. 2). Im Alter < 120 Jahre ist die ÜZW für Fichten auf Versfl. BW etwas höher, wobei sich dieser Effekt im Alter > 120 Jahre umkehrt. Im Referenzalter 100 Jahre unterscheiden sich die ÜZW für beide Datensätze nur marginal (Level I & II: 0,42; Versfl. BW: 0,45).

Der Vergleich der Ausfallgründe zwischen beiden Datensätzen ist grundlegend für die Prädiktorenselktion. Abb. 3 zeigt, dass die Fichte in beiden Datensätzen nahezu gleich häufig durch biotische Stressoren und Sturm ausfiel. Lediglich Schneebruch kam im Betrachtungszeitraum häufiger bei den auf den Versfl. BW wachsenden Fichten vor. Im Zuge der Modellierungen wurde jedoch nicht fein zwischen den Ausfallursachen unterschieden, sondern lediglich zwischen

störungsbedingter und dichtebedingter Mortalität differenziert.

Schritt II:
Prüfen der Modellgüte

Zur Ermittlung der Modellgüte wurde die ÜZW der Fichte auf den Versfl. BW mittels dem von der LWF entwickelten Modell vorhergesagt und den beobachteten Ausfallraten gegenübergestellt.

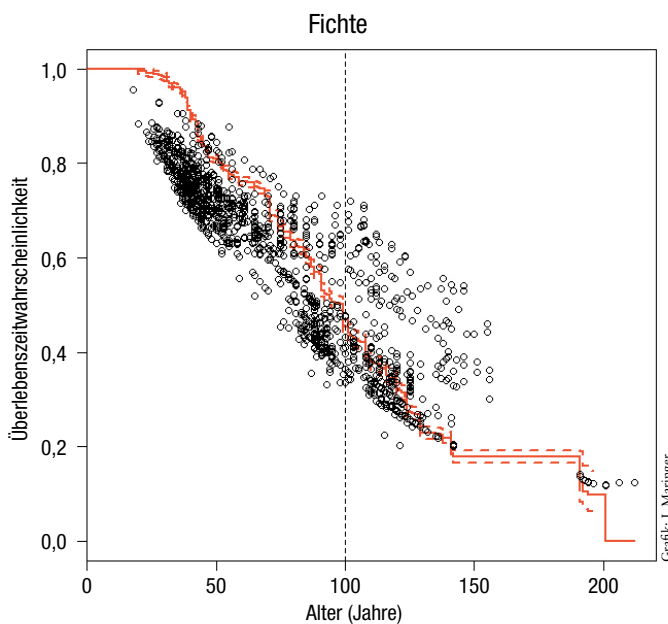


Abb. 4: Beobachtete Überlebenswahrscheinlichkeit der Fichte auf BW-Versuchsflächen (rote Kaplan-Meier-Kurve) und die Vorhersagen, basierend auf dem in Modul I entwickelten Modell (schwarze Punkte)

Der erste visuelle Eindruck (Abb. 4) zeigt, dass die mit dem Modell vorhergesagten ÜZW (schwarze Punkte) mit jenen der beobachteten ÜZW (rote Kaplan-Meier-Kurve) gut übereinstimmen. Dies wird quantitativ durch einen niedrigen Root Mean Square Error (RMSE) von 0,13 bestätigt. Der RMSE ist ein Maß zur Beurteilung der Modellgüte und gibt an, wie stark die Vorhersage im Durchschnitt von den tatsächlichen Beobachtungen abweicht. Berechnet wird der RMSE aus der Quadratwurzel des durchschnittlichen Prognosefehlers.

Schritt III: Prüfung der Modellstruktur

Der Einfluss der Prädiktoren auf die ÜZW kann in Abhängigkeit des Alters variieren. Deshalb wurden zeitliche Trends als Hinweis für die Implementierung zeitveränderlicher Variablen im Modell mittels Kaplan-Meier-Kurven überprüft. Da numerische Variablen nicht mit Kaplan-Meier-Kurven darstellbar sind, wurden die Variablen am Median in zwei Bereiche getrennt. Dargestellt werden kann somit die ÜZW

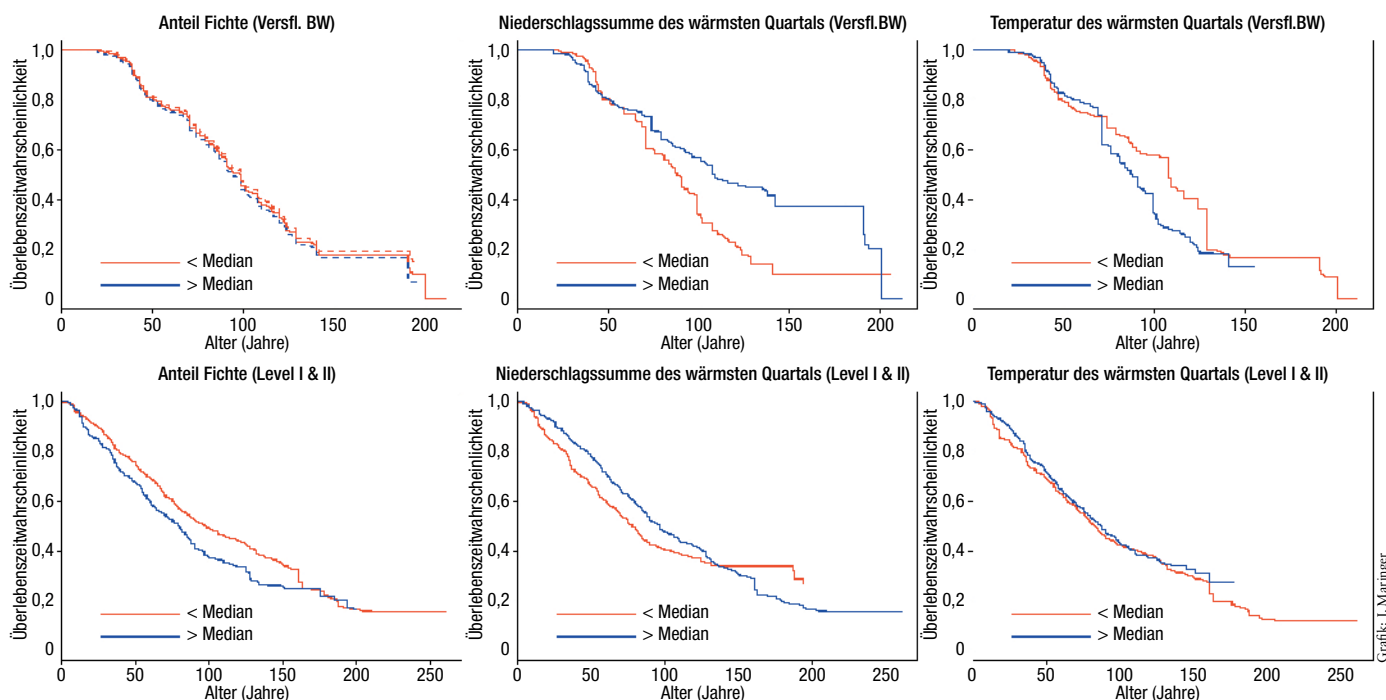


Abb. 5: Kaplan-Meier-Kurven für Fichtenanteil, Niederschlagssummen des wärmsten Quartals und Temperatur des wärmsten Quartals von Versuchsflächen Baden-Württemberg (Versfl. BW) und den Level I- & II-Daten im Vergleich

im Wertebereich unter- und oberhalb des Medians. Abb. 5 zeigt den Einfluss (ablesbar aus dem Abstand beider Kurven) sowie mögliche zeitliche Trends (Überschneidung der beiden Kurven und den damit einhergehenden Wechsel in der ÜZW). Der Fichtenanteil im Bestand zeigt weder bei den Versfl. BW noch bei den Level I- & II-Daten einen zeitlichen Trend. Erkennbar durch den Abstand zwischen den Kurven ist ein deutlich stärkerer Einfluss des Prädiktors in den Level I- & II-Daten als auf den Versfl. BW.

Die Niederschlagssummen des wärmsten Quartals zeigen eine etwas höhere ÜZW bei Werten unterhalb des Medians bis zum Alter 50 bei Fichten auf den Versfl. BW. Auf den Level I- & II-Flächen ist die ÜZW bis zum Alter 150 höher bei hohen Niederschlagssummen. Der Effekt dreht sich ab dem Alter 150 um. Bei den Temperaturen des wärmsten Quartals ist eine veränderte ÜZW in Abhängigkeit des Alters erkennbar (höhere ÜZW bei Werten oberhalb des Medians bis zum Alter 75 – danach dreht sich der Effekt um). Für diesen Prädiktor ist kein zeitlicher Trend in den Level I- & II-Daten erkennbar.

Schritt IV: Vorhersageungenauigkeiten

Zur Überprüfung der Vorhersageungenauigkeit (Richtung und Einflussgröße der Prädiktoren) wurden zwei Modelle mit standardisierten Variablen unter Verwendung des Evaluierungs- und Entwicklungsdatensatzes berechnet. Die Ergebnisse, dargestellt in Tab. 1, zeigen die gleiche Einflussrichtung für die Variablen (i) Fichtenanteil im Bestand (Pabie.ant) und (ii) Niederschlagssumme des wärmsten Quartals (Psum.wq). In Bezug auf die durchschnittliche Temperatur des wärmsten Quartals (Tmn.wq) ist ein konträrer Effekt erkennbar. Dies könnte darauf hindeuten, dass die Variable als zeitveränderliche Größe im Modell verwendet werden sollte.

Zusammenfassung und Ausblicke

Mittels einfacher Methoden konnte gezeigt werden, wie eine Modellevaluierung von AFT-Modellen, basierend auf externen Daten, erfolgen kann. Das Feld der Evaluierung von AFT-Modellen ist

Variable	Modell Level I&II	Modell Versfl. BW
	Exp(Coef)	Exp(Coef)
Pabie.ant	0,84	0,98
Psum.wq	1,037	1,165
Tmn.wq	0,898	1,172

Tab. 1: Vergleich der Einflussstärke und Wirkungsrichtung der Prädiktoren, basierend auf AFT-Modellen mit standardisierten Werten. Die Koeffizienten wurden exponenziert; Werte > 1 zeigen eine Verkürzung der Überlebenszeitwahrscheinlichkeit an, Werte < 1 das Gegenteil. Abkürzungen: Pabie.ant = Fichtenanteil im Bestand, Psum.wq: Niederschlagssummen im wärmsten Quartal, Tmn.wq: durchschnittliche Temperaturen im wärmsten Quartal (Daten basieren auf WorldClim; [12]).

aufgrund fehlender statistischer Methoden jedoch noch nicht ausreichend erschlossen. In der Risikomodellierung mit binären Antwortvariablen gibt es ausreichend etablierte Konzepte für die Modellevaluierung (vgl. [10]). Sie basiert auf der Analyse der generellen Performance (Differenz zwischen beobachteten und modellierten Werten), der Diskriminanz (Fähigkeit, Objekte mit hohem/niedrigem Risiko zu unterscheiden) und der Kalibrierung (Übereinstimmung modellierter und beobachteter Ereignisse) [13]. Die Evaluierung von Überlebenszeitmodellen ist hingegen schwieriger, da zensierte Daten den Prozess erschweren [14, 15]. Eine weitere Hürde stellen dabei linkstrunkierte und/oder intervallszensierte Daten dar [16]. Da Überlebenszeitmodelle durch ihre zensierte Datenstruktur einen Ansatz mit hoher Flexibilität in der Risikomodellierung von Baumarten bilden, ist es unabdingbar, statistische Methoden zur Modellevaluierung weiter zu entwickeln.

Das exemplarisch evaluierte Modell für die Fichte lieferte bei Anwendung auf einen externen Datensatz generell gute Modellperformanz und realistische Größenordnungen bei der Vorhersage von Mortalitätsereignissen. Mithilfe der zahlreichen Evaluierungsschritte wurden jedoch auch Defizite ersichtlich. So

Literaturhinweise:

- [1] JUTRAS, S.; HÖKÄ, H.; ALENIUS, V.; SALMINEN, H. (2003): Modelling mortality of individual trees in drained peatland sites in Finland. *Silva Fennica*. 37(2): p. 235-251. [2] HANEWINKEL, M.; PELTOLA, H.; SOARES, P.; GONZÁLEZ-OLABARRIA, J. R. (2010): Recent approaches to model the risk of storm and fire to European forests and their integration into simulation and decision support tools. *Forest Systems*. 19: p. 30-47. [3] FLEWELLING, J. W. M.; MONSERUD, R. A. (2002): Comparing methods for modelling tree mortality. in *Second Forest Vegetation Simulator Conference*. Ogden (USA): U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. [4] RICH, R. L. F.; FRELICH, L. E.; REICH, P. B. (2007): Wind-throw mortality in the southern boreal forest: effects of species, diameter and stand age. *Journal of Ecology*. 95: p. 1261-1273. [5] ADAME, P. D. R.; DEL RÍO, M.; CAÑELLAS, I. (2010): Modelling individual-tree mortality in Pyrenean oak (*Quercus pyrenaica* Willd.) stands. *Annals of Forest Science*. 67(8): p. 810. [6] COX, D. R. (1972): Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*. 34(2): p. 187-220. [7] MILLS, M. (2011): Introducing survival and event history analysis, London, New Delhi, California: SAGE Publications. [8] NEUNER, S. A.; ALBRECHT, A.; CULLMANN, D.; ENGELS, F.; GRIESS, V. C.; HAHN, W. A.; HANEWINKEL, M.; HÄRTL, F.; KÖLLING, C.; STAUPENDAHL, K.; KNOKE, T. (2014): Survival of Norway spruce remains higher in mixed stands under dryer and warmer climate. *Global Ecology & Biogeography*. [9] NEUMANN, M.; MUES, V.; MORENO, A.; HASENAUER, H.; SEIDL, R. (2017): Climate variability drives recent tree mortality in Europe. *Global Change Biology*. 23(11): p. 4788-4797. [10] BENNET, N. D.; COKE, B. F. W.; GUARISO, G.; GUILLAUME, J. H. A.; HAMILTON, S. H.; JAKEMAN, A. J.; MARSILI-LIBELLI, S.; NEWHAM, L. T. H.; NORTON, J. P.; PERRIN, C.; PIECE, S. A.; ROBSON, B.; SEPPELT, R.; VOINOV, A. A.; FATH, B. D.; ANDREASSIAN, V. (2013): Characterising performance of environmental models. *Environmental Modelling & Software*. 40: p. 1-20. [11] RAHMAN, S.; AMBLER, G.; CHOODARI-OSKOOEI, B.; OMAR, R. Z. (2017): Review and evaluation of performance measures for survival prediction models in external validation settings. *MB-MC Medical Research Methodology*. 17: p. 1-15. [12] FICK, S. E.; HIJMANS, R. J. (2017): WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*. [13] STEYERBERG, E. W. (2009): *Clinal prediction models: A practical approach to development, validation and updating*. New York: Springer. [14] CAMPBELL, H. E.; GRAY, A. M.; HARRIS, A. L.; BRIGGS, A. H.; TAYLOR, M. A. (2010): Estimation and external validation of a new prognostic model for predicting recurrence-free survival for early breast cancer patients in the UK. *British Journal of Cancer*. 103: p. 776-786. [15] ROYSTON, P.; ALTMANN, D. G. (2013): External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol*, p. 13-33. [16] BROSTRÖM, G., email from 13.06.2018 regarding residuals in aft-models using eha-package. 2018.

sollten verlässliche Mortalitätsmodelle in Zukunft beispielsweise zeitveränderliche Kovariablen, wie sich ändernde Witterungs- und Klimabedingungen, berücksichtigen.

Dr. Ing. Janet Maringer

arbeitete als wissenschaftliche Mitarbeiterin an der Forstlichen Versuchsanstalt Baden-Württemberg (FVA) und war Bearbeiterin für das Teilmodul II „Evaluierung von AFT-Modellen mittels externer Daten“ im SURVIVAL-KW-Projekt des Waldklimafonds.

Dr. Axel Albrecht

axel.albrecht@forst.bwl.de, ist ebenfalls wissenschaftlicher Mitarbeiter an der FVA und leitet das Teilmodul.

