

Mortalitätsschätzungen in ungleichaltrigen Fichtenwäldern mit Hilfe Neuronaler Netze

Hubert Hasenauer¹ und Dieter Merkl²

Abstract

Within forest growth modeling it is understood that individual tree mortality can be captured realistically by relating the average rate of mortality to a few reliable and measurable size or site characteristics using a LOGIT model. In this paper we describe the application of neuronal networks adhering to the unsupervised learning paradigm to predict individual tree mortality. Using the large and representative Norway spruce data sample from the Austrian National Forest Inventory, we train different types of neural network architectures, namely *Multi-Layer Perceptron*, *Cascade Correlation*, and *Learning Vector Quantization*. For training, we use the following learning rules: *Error Back-propagation*, *Resilient Propagation*, and *Scaled Conjugate Gradient*. With an independent data set we evaluate the neural network types to predict individual tree mortality.

1 Einleitung

Die Schätzung der Sterbe- bzw. Überlebenswahrscheinlichkeit von Bäumen als Folge nachbarschaftlicher Konkurrenz ist eine wichtige und derzeit noch schwierige Frage in der Waldwachstumsprognose. Nur wenn Bäume in einem Bestand absterben, wird das Wachstum der anderen Bäume ermöglicht. Ein wichtiger Aspekt in der Abschätzung von Baum mortalitäten ist, daß ein Baum prinzipiell nur zwei Zustände annehmen kann - tot oder lebend.

Allgemeines Ziel ist es daher herauszufinden, welche Einflußfaktoren für die Abschätzung der Sterbewahrscheinlichkeit geeignet sind. Dabei gilt es zu beachten, daß ein sehr wesentlicher Unterschied zwischen der Schätzung der Sterbe- bzw. Überlebenswahrscheinlichkeit von Bäumen im konkreten Simulationsablauf besteht, da die Wirkung möglicher Fehlschätzungen anders zu bewerten ist. Beispielsweise kann ein Baum, der irrtümlich als lebend eingestuft wird, im nächsten Prognosezeitraum sehr wohl absterben, während der umgekehrte Fall, ein Baum wird als abgestorben

¹ Universität für Bodenkultur Wien, Institut für Waldwachstumsforschung, Peter-Jordan-Str. 82, A-1190 Wien. email: hasenau@edv1.boku.ac.at

² Institut für Softwaretechnik, Technische Universität Wien, Resselgasse 3/188, A-1040 Wien. Internet: <http://www.ifs.tuwien.ac.at/~dieter>

klassifiziert, sollte aber überleben, nicht mehr möglich ist, weil tote Bäume aus der Baumliste eliminiert werden. Bedenkt man weiters, daß die jährlichen Mortalitätsraten unserer mitteleuropäischen Wälder je nach Altersphase und Baumart 1 bis 10 % betragen, so wird deutlich, daß eine genauere Mortalitätsschätzung eine wesentliche Verbesserung der Waldwachstumssimulation bedeuten würde (Hasenauer 1999).

Das Forschungsgebiet der künstlichen Neuronalen Netze konnte sich in den letzten Jahren für Bereiche, die analytische oder algorithmische Lösungsansätze nicht ermöglichen, als außerordentlich erfolgreiches Feld der Informatik etablieren. Das gilt insbesondere für Anwendungsgebiete mit einer großen Anzahl an Einflußgrößen, die zusätzlich noch unbekannte Zusammenhänge aufweisen. Ist für ein solches Anwendungsgebiet jedoch eine genügend große Zahl an Trainingsbeispielen verfügbar, die die mögliche Datenkonstellationen gut abdecken, so ist eine erfolgreiche Problemlösung mit Hilfe künstlicher Neuronaler Netze zu erwarten.

Neuronale Netze werden anhand von Beispielen trainiert. Eine Problemlösung wird sozusagen erlernt und nicht, wie in konventionellen Modellen der Informatik mit Hilfe eines Programms realisiert. Aus der Literatur sind eine Reihe möglicher Aufgabenstellungen, die mit Hilfe künstlicher Neuronaler Netze erfolgreich gelöst werden können, bekannt (Swingler 1996).

In unseren bisherigen Arbeiten zur Verwendung Neuronaler Netze in der Abschätzung von Baummortalitäten haben wir uns mit der Analyse von *Self-Organizing Maps* (Hasenauer/Merkl 1997) und *Learning Vector Quantization* (Merkl/ Hasenauer 1998) an einem kleinen Datensatz beschäftigt. Dabei hat sich gezeigt, daß die Schätzung der Baummortalität mit Hilfe Neuronaler Netze stabiler als der konventionelle statistische Ansatz der LOGIT Modelle ist.

Basierend auf diesen Ergebnissen ist das Ziel dieser Arbeit die Anwendung verschiedener neuronaler Netztypen an einem wesentlich umfangreicheren und vor allem repräsentativ für ganz Österreich erhobenen Datensatz der Baumart Fichte (*Picea abies* L. Karst) zu vergleichen. Dieser Datensatz beinhaltet alle in Österreich vorkommenden Bestandestypen der Fichte und stellt daher eine ausgezeichnete Datengrundlage für eine Analyse verschiedenster Netztypen und deren Eignung für die Nachbildung von Baummortalitäten dar.

Konkret werden folgende Architekturen Neuronaler Netze in dieser Studie getestet: (1) *Multilayer Perceptron* (Bishop 1995), (2) *Cascade Correlation* (Fahlman/Lebiere 1990), (3) *Learning Vector Quantization* (Kohonen 1990). Bei der Auswahl der Netztypen wurde besonders darauf geachtet, daß drei fundamental unterschiedliche überwachte Modelle zum Einsatz gelangen. Weiters wurde mit dem *Multilayer Perceptron* quasi der Archetyp des künstlichen Neuronalen Netzes auf seine Anwendbarkeit für Waldwachstumsmodellierung untersucht.

Datengrundlage für unsere Studie sind die Daten der Österreichischen Waldinventur (Forstliche Bundesversuchsanstalt 1994), die auf ca. 5.500 Traktpunkten Wachstums- und Mortalitätsinformation von über 43.000 Bäumen gesammelt hat. Diese Daten waren auch die Grundlage für das derzeit auf Basis des konventionellen

statistischen Ansatzes in Verwendung befindliche Mortalitätsmodell von Monserud/Sterba (1999). Für die Evaluierung der Schätzergebnisse steht ein unabhängiger Datensatz, der somit einen konkreten Anwendungsfall für die trainierten Netze repräsentiert, zur Verfügung.

2 Neuronale Netze

Die derzeitigen Lernverfahren für Neuronale Netze repräsentieren sehr robuste Methoden zur Approximation von Zielfunktionen. Insbesondere hat sich in einer großen Vielzahl von Anwendungsgebieten die Toleranz dieser Lernverfahren gegenüber verrauschten oder unvollständigen Eingabedaten erwiesen.

Grundsätzlich bestehen Neuronale Netze aus einer Menge von sogenannten Neuronen, die untereinander über gewichtete Verbindungen kommunizieren. Dabei besteht die Kommunikation im Empfangen und Versenden von Impulsen. Ein einzelnes Neuron empfängt typischerweise von vielen Neuronen Impulse, berechnet aufgrund der verwendeten Aktivierungsfunktion seine eigene Aktivierung, und schickt in Abhängigkeit von dieser Aktivierung einen Impuls an ebenso eine Vielzahl von Neuronen.

Bezüglich der verwendeten Lernverfahren werden Neuronale Netze für gewöhnlich in (1) überwachte, (2) bestärkende und (3) unüberwachte Modelle eingeteilt. Diese Modelle unterscheiden sich hinsichtlich der Information über das erwünschte Trainingsergebnis, das dem Neuronalen Netz während des Trainings zur Verfügung gestellt wird (Anderson 1995). In weiterer Folge werden wir im Rahmen dieser Arbeit ausschließlich Neuronale Netze mit überwachter Lernfunktion betrachten.

2.1 Multi-Layer Perceptron (MLP)

Die wohl bekannteste Architektur Neuronaler Netze ist das sogenannte *Multi-Layer Perceptron*. Diese Neuronalen Netze bestehen aus Neuronen, die jeweils in Schichten angeordnet sind. Neuronen einer Schicht haben typischerweise ausschließlich Verbindungen zu Neuronen der unmittelbar benachbarten Schichten. Eine Schicht wird als Eingabeschicht ausgewiesen und erhält seine Eingaben nicht von anderen Neuronen sondern von außen (d.h. von einem umgebenden System, in das Neuronale Netz eingebettet ist). Analog dazu wird eine Schicht als Ausgabeschicht verwendet, die ihre Ausgaben nicht an andere Neuronen sondern nach außen abgibt. Allen übrigen Schichten werden als verdeckte Schichten bezeichnet. Die Verbindungen zwischen Neuronen sind gewichtet. Mit w_{ij} bezeichnen wir in weitere Folge das Gewicht der Verbindung von Neuron i zu Neuron j .

2.2 Cascade Correlation (CC)

Eines der Probleme mit dem *Multi-Layer Perceptron* ist, daß jeweils sämtliche Neuronen teilhaben an der Abbildung eines Eingangswertes. Das fällt insbesondere dann ins Gewicht, wenn das Neuronale Netz bereits eine Reihe von Eingangsdaten korrekt behandelt, durch die Präsentation eines neuen Wertes jedoch die Netzwerkkonfiguration derart verändert wird, daß zuvor korrekt behandelte Eingangsdaten nunmehr fehlerhaft abgebildet werden. Diesem Problem soll die *Cascade Correlation* Architektur derart begegnen, daß nicht nur die Gewichte von Verbindungen während des Trainings verändert werden, sondern auch der Aufbau des Netzwerkes durch wiederholtes Hinzufügen von Neuronen (Fahlman/Lebiere 1990).

Zu Beginn des Trainings existieren nur die durch die Anwendung vorgegebenen Eingabe- bzw. Ausgabeneuronen. Diese sind über gewichtete Verbindungen verbunden. Das Lernverfahren fügt nun einzeln verdeckte Neuronen zu dem Netzwerk hinzu. Dabei erhält jedes neue Neuron Eingaben von allem bis dahin existierenden Neuronen, d.h. von den Eingabeneuronen und bereits eingefügten verdeckten Neuronen. Jedes neue Neuron erhält weiters gewichtete Verbindungen zu allen Ausgabeneuronen.

Bei diesem Neuronalen Netz bleiben jeweils nur die Gewichte von Verbindungen zur Ausgabeschicht durch die Lernregel veränderbar. Die Gewichte zu bereits "älteren" Neuronen werden nicht mehr verändert. Diese Strategie führt dazu, daß mit den neu hinzugefügten Neuronen genau die durch die davor bestehende Architektur nicht ausreichend reduzierbare Abweichung zwischen berechnetem und gewünschtem Ergebnis verringert wird.

2.3 Überwachte Lernverfahren

In diesem Abschnitt sollen jene Lernverfahren für überwachte Neuronale Netze kurz vorgestellt werden, die in weiterer Folge in dieser Arbeit Anwendung finden. Aus Platzgründen kann diese Übersicht jedoch nur äußerst knapp ausfallen. Wir ersuchen daher den interessierten Leser/die interessierte Leserin auf Arbeiten wie Anderson (1995), Bishop (1995), Kohonen (1988) oder Zell (1994) zurückzugreifen. Eine stark anwendungsorientierte Darstellung der Lernverfahren findet sich im User Manual des SNNS³ (Zell *et al.* 1995).

³ Das Kürzel SNNS steht für *Stuttgart Neural Network Simulator*, ein Werkzeug zum Aufbau und Training von Neuronalen Netzen. SNNS kann über den FTP-Server der Universität Stuttgart bezogen werden: <ftp.informatik.uni-stuttgart.de/pub/SNNS>.

2.3.1 Error Backpropagation (BP)

Das wohl bekannteste Lernverfahren für überwachte Neuronale Netze ist *Error Backpropagation* (Rumelhart et al. 1986). Diese Lernregel beschreibt ein Gradientenabstiegsverfahren, wo jeweils in Richtung des steilsten Abstiegs der Fehlerfunktion des Neuronalen Netzes versucht wird, ein globales Minimum zu finden. Mit anderen Worten, es wird nach jener Konfiguration der Gewichte gesucht, so daß der Fehler des Neuronalen Netzes (d.h. die Differenz zwischen errechnetem und gewünschtem Ergebnis) minimiert wird.

Mit o_i als Ausgabe von Neuron i , δ_j als gewichteter Fehler von Neuron j (berechnet als Differenz zwischen gewünschtem und errechnetem Ergebnis) und a als Lernrate (Parameter zur Steuerung der Größe der Gewichtsveränderung können wir die Veränderung Δw_{ij} des Gewichtes der Verbindung von i nach j aufgrund von Trainingsschritt t wie in der folgenden Gleichung beschreiben.

$$\Delta w_{ij}(t) = \alpha(t) o_i(t) \delta_j(t)$$

Zur Herleitung des gewichteten Fehlers δ_j wird dabei die erste Ableitung der Aktivierungsfunktion gemeinsam mit der Abweichung zwischen errechnetem und gewünschtem Ergebnis verwendet. Dies geht unmittelbar für die Gewichte zwischen der letzten verdeckten Schicht und der Ausgabeschicht. Die übrigen gewichteten Fehler können nur indirekt berechnet werden, wobei jeweils ein Teil des Fehlers des dahinterliegenden Neurons reduziert wird. Wir verzichten auf die mathematische Darstellung der Berechnung, sie kann in jedem Lehrbuch der Neuronalen Netze nachgeschlagen werden, so z.B. in (Zell 1994).

2.3.2 Resilient Propagation (Rprop)

Im vorigen Abschnitt haben wir *Error Backpropagation* als Gradientenabstiegsverfahren beschrieben. Die Größe der Gewichtsveränderungen wurde dabei von einer Lernrate bestimmt. Diese war jedoch unabhängig von der konkreten Form der Fehlerkurve.

Eine Beschleunigung des Lernverfahrens ist möglich, wenn man die konkreten Gegebenheiten der Fehlerfunktion in Betracht zieht: Größere Schrittweite bei gleichbleibendem Gradienten, kleinere Schrittweite bei Richtungsänderung des Gradienten. Das von Riedmiller/Braun (1993) vorgestellte Lernverfahren für mehrschichtige, überwachte Neuronale Netze, Rprop, beruht genau auf dieser Beobachtung. Die Schrittweite wird dabei durch das Vorzeichen des Gradienten bestimmt.

2.3.3 Scaled Conjugate Gradient (SCG)

Wurde in den bisher beschriebenen Lernverfahren jeweils die erste Ableitung der Aktivierungsfunktion verwendet, so beruht bei *Scaled Conjugate Gradient* (Moller 1993) die Gewichtsveränderung auf der zweiten Ableitung. Ein solches Lernverfahren konvergiert erfahrungsgemäß schneller in Richtung eines (lokalen) Minimums als Gradientenabstiegsverfahren. Die Kehrseite der Medaille ist jedoch der deutlich höhere Berechnungsaufwand generell für *Conjugate Gradient* Verfahren.

3 Daten

3.1 Trainingsdaten

Die Daten für das Training der unterschiedlichen Netztypen stammen von der Österreichischen Waldinventur (Forstliche Bundesversuchsanstalt 1994). Die Inventur basiert auf einem systematisch über ganz Österreich verteilten Traktsystem im Abstand von 3,89 km und einem 5-jährigen Aufnahmeintervall. Jeder Trakt besteht aus einem Quadrat von 200 x 200 m Seitenlänge, in dessen Ecken die permanenten Probepunkte eingerichtet wurden. Jährlich wurde ein fünftel der Trakte aufgemessen. Somit wurde sichergestellt, daß jedes Jahr eine repräsentative Aufnahme für ganz Österreich erfolgte.

Alle Probepunkte sind so markiert, daß sie zwar auffindbar, aber nicht im Wald sichtbar sind. Auf diese Weise wird die Repräsentativität der Daten für das Waldwachstum und die Waldbehandlung in Österreich sichergestellt. Die gesamte Inventur besteht aus 5.500 Trakten bzw. 22.000 Probepunkten. Knapp weniger als die Hälfte der Punkte befinden sich im österreichischen Wald.

Die permanenten Probepunkte wurden zwischen 1981 und 1985 eingerichtet. Bäume mit einem Brusthöhendurchmesser *DBH* von größer als 10,4 cm wurden mittels Winkelzählprobe (Bitterlich 1948) mit einem Zählerfaktor von 4 m²/ha ausgewählt. Bäume mit einem *DBH* zwischen 5 und 10,4 cm wurden nur dann erhoben, wenn sie innerhalb des fixen Probekreises von 2,6 m Radius rund um das Probeflächenzentrum standen. Bäume mit einem *DBH* von weniger als 5 cm wurden nicht erhoben. Bei der Erstaufnahme wurden für jeden Baum am Probepunkt folgende Erhebungen durchgeführt: Baumart, Brusthöhendurchmesser auf mm, die Distanz und der Azimut vom Probeflächenzentrum, Baumhöhe und Kronenansatzhöhe auf dm. Die Bestandesdichtemaße *BAL* (Grundfläche der stärkeren Bäume) und *CCF* (Crown Competition Factor) wurden für jeden Aufnahmepunkt berechnet.

Sämtliche Bäume auf den Probepunkten wurden zwischen 1986 und 1990, also genau nach 5 Jahren, wieder aufgemessen. Dabei wurde von jedem Baum der Brusthöhendurchmesser gemessen, aber nur von jedem fünften Baum auf der Fläche die Baumhöhe. Der Kronenansatz wurde nicht mehr gemessen. Insgesamt stehen 26.699 Beobachtungen an Fichtenbäumen verteilt über 4.135 Probepunkte in Österreich für

diese Studie zur Verfügung. Die mittlere Sterblichkeit dieses Datensatzes beträgt 0,9 % pro Jahr. Mit diesen Daten wurden eine Reihe von Neuronalen Netzen trainiert. In einem Vorverarbeitungsschritt wurden die Einträge für *DBH* und *BAL* linear auf das Intervall [0, 1] normiert. Mittels Cross-Validierung werden die besten Neuronalen Netze ausgewählt.

3.2 Evaluierungsdaten

Im Jahre 1977 wurden 22 permanente Probeflächen unterschiedlichen Alters und verschiedener Bonität in gemischten ungleichaltrigen Fichten-Kiefernbeständen der Seilern-Aspang'schen Forstverwaltung in Litschau eingerichtet. Der Forstbetrieb liegt in Niederösterreich, der nördlichsten Ecke des Waldviertels.

Der vorliegende Bestandestyp der Fichten-Kiefermischbestände entstand als Folge des Kahlschlagbetriebes mit nachfolgender natürlicher Ansamung, die zum Teil mit Fichten ergänzt wurde. Die Probeflächengröße variierte je nach Alter bei der Ersteinrichtung zwischen 400 m² für junge, sehr stammzahlreiche Bestände und 1600 m² für die Altbestände. Jede der Probeflächen wurde in den Jahren 1982, 1987, 1992 und 1997 neu aufgemessen. Dabei wurden für jeden Baum auf der Probefläche die Baumhöhe, der Brusthöhendurchmesser und die Kronenansatzhöhe in den jeweiligen Aufnahmejahren erhoben.

Im Zuge der Wiederholungsaufnahmen wurden sämtliche Bäume, die eine Baumhöhe von 1,3 m (Brusthöhe) erreicht hatten, als Einwüchse mitaufgenommen und in der Interpretation entsprechend berücksichtigt. Um das Wachstum in Abhängigkeit von den konkurrenzbeschreibenden Faktoren studieren zu können, sind Unterschiede in der Bestandesdichte auf den Probeflächen eine der Grundvoraussetzungen. Im Laufe der ersten Wachstumsperiode (1977-1982) verursachte ein Schneebruch in einigen Parzellen eine Stammzahlabnahme, die nicht konkurrenzbedingt erklärbar ist und ein Katastropheneignis darstellt. Die Daten dieser Wachstumsperiode wurde daher nicht für die Evaluierung berücksichtigt.

3 Analyse und Ergebnisse

Für die Schätzung der Mortalitätswahrscheinlichkeiten am Beispiel der Daten der Österreichischen Waldinventur entwickelten Monserud/Sterba (1999) ein Mortalitätsmodell auf Basis des LOGIT Ansatzes (Neter/Maynes 1970) folgender Form:

$$P = \frac{1}{1 + e^{(b_0 + b_1 / DBH + b_2 CR + b_3 BAL + b_4 DBH + b_5 DBH^2)}}$$

P ist die Mortalitätswahrscheinlichkeit eines Baumes in einer 5-Jahresperiode, DBH der Brusthöhendurchmesser, CR das Kronenverhältnis und BAL die Kreisflächendichte der stärkeren Bäume.

Ähnliche Modelle stellen das heute übliche Verfahren für die Abschätzung der Mortalitätsraten dar und bilden somit einen hervorragenden Ausgangspunkt für die Anwendbarkeit bzw. den Vergleich von alternativen Ansätzen. In unserer Studie werden wir daher die mit Hilfe Neuronaler Netze erzielten Ergebnisse jeweils mit dem von Monserud/Sterba (1999) auf Basis des klassischen LOGIT Ansatzes berechneten Ergebnissen kontrastieren.

Für die weitere Vorgehensweise und die Interpretation der Ergebnisse ist wichtig, daß der Vektor der Variablen für die Beschreibung der individuellen Mortalität durch das LOGIT Modell vorgegeben ist. Wir verwenden also keine zusätzlichen Variablen, da ansonsten ein direkter Vergleich von LOGIT-Schätzung und Schätzung auf Basis Neuronaler Netze erschwert würde.

Abbildung 1 zeigt die beobachteten Mortalitätsraten (Real) im Vergleich zum LOGIT Ansatz und folgenden getesteten Neuronalen Netztypen bzw. verwendeten Lernregeln: *Multi Layer Perceptron* (MLP) auf Basis von *Error Backpropagation* – MLP(BP); MLP mit Hilfe *Resilient Propagation* MLP(Rprop); MLP mit Hilfe *Scaled Conjugate Gradient* MLP(SCG); *Cascade Correlation* (CC) auf Basis des *Error Backpropagation* - CC(BP); CC mit *Resilient Propagation* – CC(Rprop) und das *Learning Vector Quantization* - LVQ. Aus Abbildung 1 ist ersichtlich, daß das LVQ Netz die schlechtesten Schätzergebnisse erzielt hat. In allen folgenden Analysen wird daher auf diesen Netztyp verzichtet.

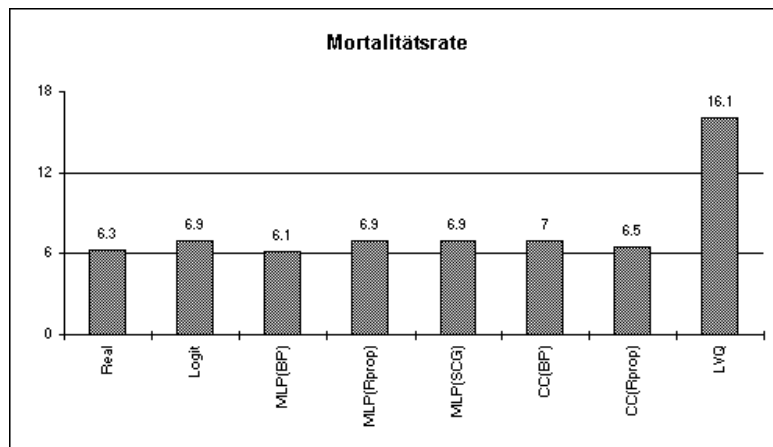


Abbildung 1

Gegenüberstellung der beobachteten (Real) und je Ansatz geschätzten Mortalitätsraten in Prozent der Gesamtbaumzahl für den unabhängigen Datensatz aus Litschau

Als nächstes waren die absoluten Abweichungen der Schätzungen von den auf den Probestflächen beobachteten Mortalitätsraten von Interesse. Abbildung 2 zeigt die Ergebnisse der einzelnen Verfahren je nach untersuchter Lernregel. In dieser Grafik wird deutlich, daß das *Multi-Layer Perceptron* mit *Error Backpropagation* Lernregel, MLP(BP), und *Cascade Correlation* mit *Resilient Propagation*, CC(Rprop), wesentlich bessere Ergebnisse liefern als das LOGIT Modell. Die übrigen Neuronalen Netze schätzen unwesentlich bessere oder gleich gut als der konventionell statistische LOGIT Ansatz.

Die Mortalitätsraten von Waldbeständen hängen sehr wesentlich vom Bestandesalter ab. Die für die Evaluierung herangezogenen Einzelbaumdaten der 23 Probestflächen können als typische Waldbestände für den Raum Litschau verstanden werden, die je nach Altersphase verschiedene Stadien in der Waldentwicklung und damit den zu erwartenden Mortalitätsraten widerspiegeln. Das Bestandesalter dieser Flächen variiert zwischen 10 und 110 Jahren. Um nun mögliche systematische Effekte hinsichtlich verschiedener Bestandesalter evaluieren zu können, wurden die 23 Waldbestände in 5 Altersgruppen eingeteilt. Abbildung 3 zeigt für jede Altersgruppe den relativen Fehler der prognostizierten Mortalitätsraten. Die Verteilung der Bäume im Evaluierungsdatensatz auf die einzelnen Klassen wird in Tabelle 1 dargestellt.

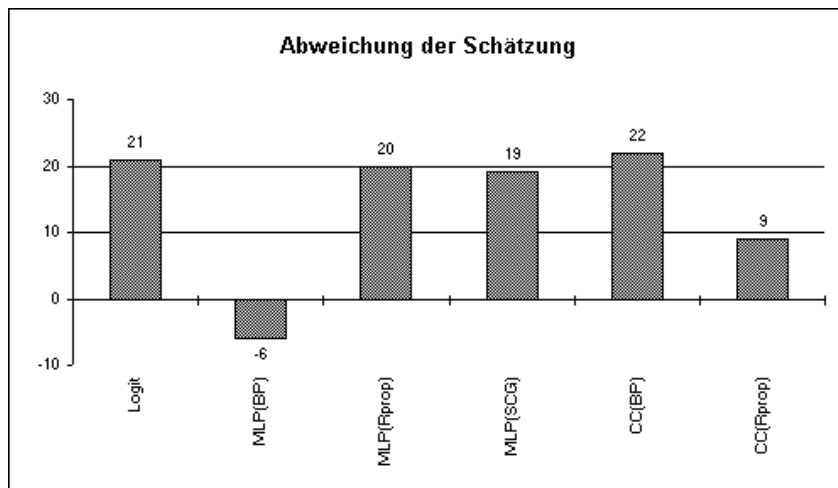


Abbildung 2

Abweichung der absoluten Anzahlen von geschätzten toten Bäumen vom wahren Wert für das LOGIT Modell und die 5 getesteten Neuronalen Netze

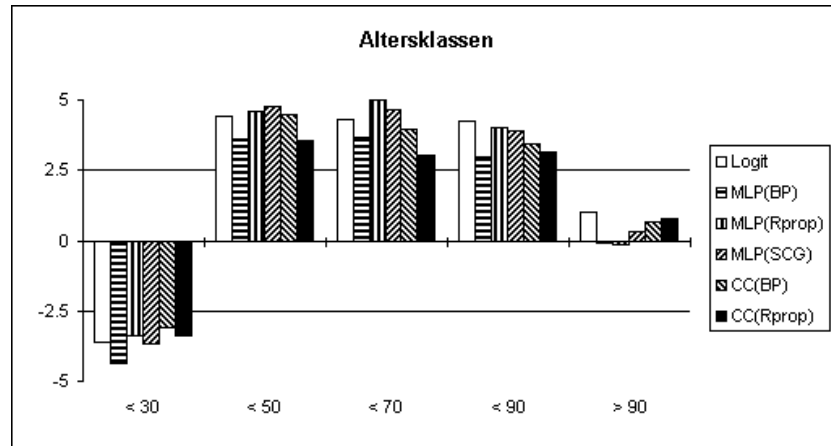


Abbildung 3

Prozentueller Fehler der geschätzten Mortalitätsraten für das LOGIT Modell und die Neuronalen Netztypen getrennt für unterschiedliche Bestandesalter der einzelnen Probeflächen

<i>Altersklasse</i>	<i>Anzahl lebende Bäume</i>	<i>Anzahl tote Bäume</i>
Alter kleiner 20 Jahre	1096	123
Alter kleiner 50 Jahre	619	47
Alter kleiner 70 Jahre	368	11
Alter kleiner 90 Jahre	316	3
Alter größer 90 Jahre	551	13

Tabelle 1

Verteilung der Evaluierungsdaten auf die einzelnen Altersklassen

In der folgenden Abbildung 4 präsentieren wir eine detailliertere Auswertung der Prognoseergebnisse bezogen auf die maßgeblichen Eingangsvariablen *DBH* (Brusthöhendurchmesser), *CR* (Kronenverhältnis) und *BAL* (Grundfläche der stärkeren Bäume). Eine derartige Analyse ist deshalb wichtig, weil Modellschätzungen im Mittel durchaus richtig erscheinen mögen aber möglicherweise Trends in der Größe des Wertes bestimmter Eingangsvariablen aufweisen könnten, die dann auf einen systematischen Fehler im Modell hinweisen würden. Die Bäume der Evaluierungsdaten werden wieder jeweils in fünf Gruppen je nach den konkreten Werten der Eingangsvariablen eingeteilt und hinsichtlich ihrer Fehlschätzungen evaluiert (siehe Abbildung 4).

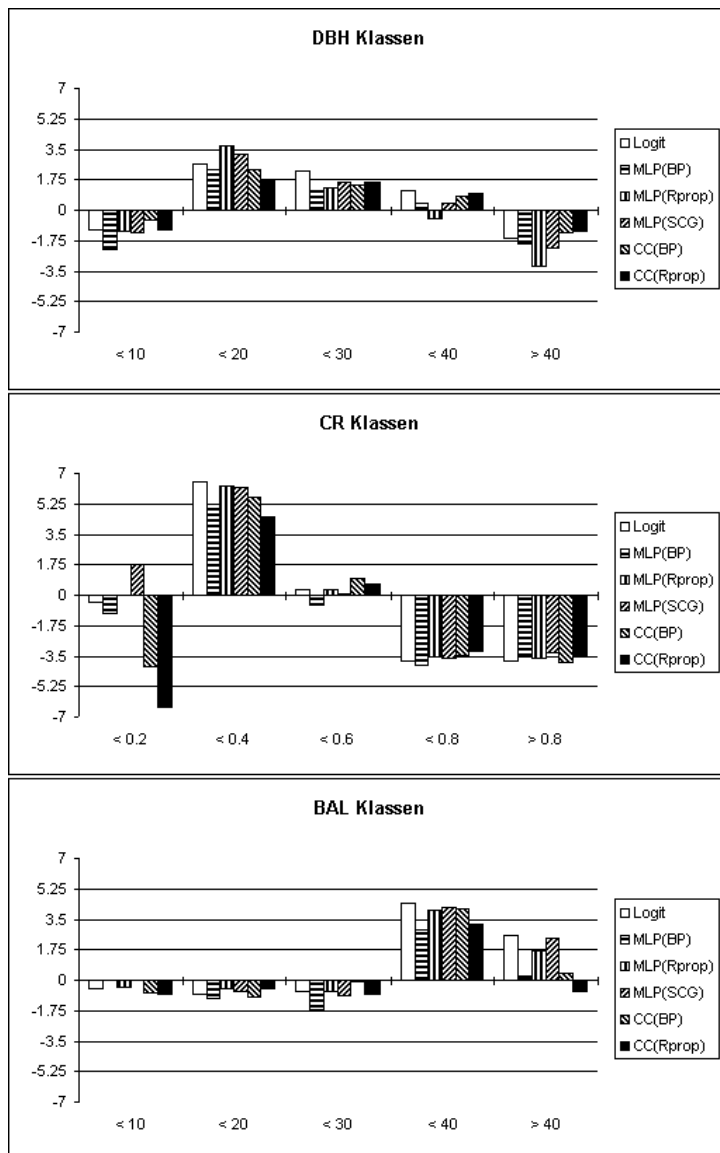


Abbildung 4
 Prozentueller Fehler der geschätzten Mortalitätsraten je
 Schätzmethode dargestellt für die wichtigsten Variablen

4 Diskussion

Ein erster Vergleich der verschiedenen Netztypen zeigt, daß die Methode des *Learning Vector Quantization* die deutlich schlechtesten Prognoseergebnisse erbracht hat (siehe Abbildung 1). Mit 16,1 % geschätztem Mortalitätsanteil bezogen auf die Gesamtbaumzahl entspricht das einer Überschätzung von 9,8 % gegenüber der tatsächlich in Litschau im Mittel beobachteten Mortalität von 6,3 %. Dieses Ergebnis steht in Widerspruch zu einer unserer früheren Studien (Merkl/Hasenauer 1998), bei dem diese Fehlschätzung auf Grund einer etwas anderen Datenlage nicht erkennbar war. Alle anderen getesteten Netztypen resultierten in wesentlich besseren Schätzergebnissen und weisen im Mittel einen relativen Unterschied zum wahren Wert von weniger als 1 % auf. Die besten Ergebnisse werden mit dem *Multi Layer Perceptron* mit Lernregel *Error Backpropagation* MLP(BP) und dem *Cascade Correlation* mit Lernregel *Resilient Propagation* CC(Rprop) erreicht (siehe Abbildung 1 und 2). Die Treffergenauigkeit ist besser als für den heute üblichen LOGIT Ansatz.

Eine wichtige Erkenntnis ergibt sich aus Abbildung 3: In jungen Beständen wird die Mortalität von allen Schätzverfahren also auch vom LOGIT Ansatz etwas unterschätzt, während mit zunehmendem Bestandesalter eine leichte Überschätzung der Mortalitätsraten zu erkennen ist. In diesem Zusammenhang ist wichtig zu beachten, welche Wirkungen dieser, wenn auch geringe (< 5%) Fehler für die praktisch Waldwachstumssimulation haben kann. Wie bereits erwähnt, hat die Unterschätzung der Mortalität eine andere Bedeutung als eine Überschätzung, da Bäume, die etwa in der Altersklasse < 30 irrtümlich für einen bestimmten 5 jährigen Prognosezeitraum als lebend eingestuft werden, in der nächsten Periode absterben können. Im umgekehrten Fall ist das nicht möglich, weil tote Bäume aus der Baumliste eliminiert werden und damit in der weiteren Waldwachstumsprognose fehlen. Basierend auf dieser Überlegung ist die leichte Unterschätzung als eher unproblematisch zu erachten. Ebenso die leichte Überschätzung in den älteren Beständen. Die Anzahl der gestorbenen Bäume ist gering (siehe Tabelle 1), daher kommt der Fehlschätzung von nur einem Baum ein wesentlich höheres relatives Gewicht zu als etwa in den jungen, sehr stammzahlreichen Beständen.

In Abbildung 4 sind nun durchaus unterschiedliche Trends in Bezug auf Unter- bzw. Überschätzung der Mortalität zu beobachten. Ein sehr schönes Beispiel dafür ist etwa die erste CR Klasse mit Bäumen, die ein sehr geringes Kronenverhältnis aufweisen. In dieser Klasse sind die Abweichungen bei den Neuronalen Netzen fast durchwegs weitaus größer als beim LOGIT Ansatz. Das schlechte Abschneiden der Neuronalen Netze in dieser Klasse ist aber insofern zu relativieren, da hier nur vergleichsweise wenige Bäume zu finden sind. Eine Fehlschätzung wirkt sich somit wiederum prozentuell sehr stark aus, hat jedoch absolut gesehen wenig Einfluß auf das Gesamtergebnis.

5 Zusammenfassung

In dieser Arbeit wurde eine Serie von Experimenten beschrieben, die die grundsätzliche Eignung Neuronaler Netze für einen wichtigen Aspekt der Waldwachstumsmodellierung getestet hat. Konkret ging es dabei um die Schätzung von Mortalität in ungleichaltrigen Wäldern. Mortalitätsschätzungen sind notwendig, um realistische Stammzahlabnahmen in Waldwachstumssimulatoren berechnen zu können. Im Rahmen dieser Studie galt unser Hauptaugenmerk dem Vergleich der Ergebnisse zwischen dem klassischen LOGIT Ansatz und unterschiedlichen Neuronalen Netzen, wobei wir uns hier auf Modelle mit überwachter Lernfunktion beschränkt haben.

Die Ergebnisse haben gezeigt, daß Neuronale Netze prinzipiell für Mortalitätschätzungen verwendbar sind. Mit Ausnahme des LVQ Netzes, ergab sich für keines der evaluierten Modelle wesentlich schlechtere Ergebnisse als der etablierte LOGIT Ansatz. Im Gegenteil, in einigen Fällen waren die Ergebnisse besser.

Wir benutzen diese ermutigenden Beobachtungen als Ausgangspunkt für weitere Experimente. Dabei werden wir jedoch die den Prognosen zugrundeliegenden Rohdaten erweitern. Insbesondere denken wir an das Einbeziehen von Daten über Boden- und Humusbeschaffenheit, Hangneigung, Exposition und Seehöhe sowie weiterer Kenngrößen für die Beschreibung der Konkurrenzsituation einzelner Bäume in einem Waldbestand.

Die gefundenen Ergebnisse sind Teil wichtiger methodischer Grundlagen für die Waldwachstumsprognose auf Basis von Einzelbaummodellen. Derartige Modelle werden entwickelt, um die Komplexität der Baumartenzusammensetzung und die mögliche Ungleichaltrigkeit unserer Wälder richtig nachbilden zu können. Nachdem diese Art der Waldwachstumsprognose die Möglichkeit schafft, eine Reihe ökologisch relevanter Parameter, wie Diversität, Struktur und Naturhöhe für unsere Wälder zu berechnen, stellt die Adaptierung Neuronaler Netze für die Mortalitätschätzungen im Rahmen der Einzelbaummodellierung einen wichtige Beitrag für die nachhaltige Bewirtschaftung und den Schutz unserer Wälder dar.

6 Danksagung

Für die zur Verfügung gestellten Daten der Österreichischen Waldinventur gilt unser Dank Herrn Hofrat Dr. Karl Schieler, Leiter des Institutes für Waldinventur an der Forstlichen Bundesversuchsanstalt Wien. Unser Dank gilt weiters Herrn Martin Weingartner für seinen Einsatz beim Training der Neuronalen Netze und den drei Begutachtern dieser Arbeit. Die Studie ist Teil des Projektes *Ökosystemmodellierung mit Neuralen Netzen*, das vom Jubiläumsfonds der Österreichischen Nationalbank (Projektnummer 7058) gefördert wird.

Literaturverzeichnis

- Anderson, J.A. (1995): *An Introduction to Neuronal Networks*, Cambridge/MA
- Bishop, C. (1995): *Neural Networks for Pattern Recognition*, Oxford, UK
- Bitterlich, W. (1948): Die Winkelzählprobe, in: *Allgemeine Forst- und Holzwirtschaftliche Zeitung*, 59, Heft 1/2, S. 4-5
- Fahlman, S.E., Lebiere, C. (1990): The cascade-correlation learning architecture, in: Touretzky, D. S. (ed.): *Advances in Neural Information Processing Systems 2*, San Mateo/CA, S. 524-532
- Forstliche Bundesversuchsanstalt (1994): *Instruktionen für die Feldarbeit der Österreichischen Forstinventur 1981-1985*, Forstl. Bundesversuchsanstalt, Wien
- Hasenauer, H., Merkl, D. (1997): Forest Tree Mortality Simulation in Uneven-Aged Stands Using Connectionist Networks, in: *Bulsari, A.B., Kallio, S. (eds.): Proceedings of the Int'l Conference on Engineering Applications of Neural Networks*, Stockholm, Sweden, pp. 341-348
- Hasenauer, H. (1999): *Die simultanen Eigenschaften von Waldwachstumsmodellen*, Berlin
- Kohonen, T. (1988): An Introduction to Neural Computing, in: *Neural Networks*, 1, pp. 3-16
- Kohonen, T. (1990): Improved versions of learning vector quantization, in: *Proceedings of the IEEE Int'l Conference on Neural Networks*, San Diego/CA, pp. 545-550
- Merkl, D., Hasenauer, H. (1998): Using Neural Networks to Predict Individual Tree Mortality, in: *Proceedings of the Int'l Conference on Engineering Applications of Neural Networks*, Gibraltar, UK, pp. 197-204
- Moller, M.F. (1993): A scaled conjugate gradient algorithm for fast supervised learning, in: *Neural Networks*, 6, pp. 525-533
- Monserud, R.A., Sterba, H. (1999): Modeling individual tree mortality for Austrian forest species, in: *Forest Ecol. and Manage.*, 113, pp. 109-123
- Neter, J., Maynes, S.E. (1970): On the appropriateness of the correlation coefficient with a 0,1 dependent variable, in: *J. Am. Stat. Assoc.*, 65, No. 350, pp. 501-509
- Riedmiller, M., Braun, H. (1993): A direct adaptive method for faster backpropagation learning: The RPROP algorithm, in: *Proceedings of the IEEE Int'l Conference on Neural Networks*, San Francisco/CA, pp. 586-591
- Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986): Learning representations by back-propagation errors, in: *Nature*, 323, pp. 533-536
- Swingler, K. (1996): *Applying Neural Networks: A Practical Guide*, London
- Zell, A. (1994): *Simulation Neuronaler Netze*, Bonn et al.
- Zell, A. et al. (1995): *SNNS - Stuttgart Neural Network Simulator, User Manual, Version 4.1*, Universität Stuttgart