

# **Das Datenbanksystem 'ECO'**

## **am Beispiel der Verwaltung standortskundlicher Erhebungen**

(mit Systemvorführung)

Andreas Schulze

Niedersächsische Forstliche Versuchsanstalt, Grätzelstr. 2, D-37079 Göttingen  
aschu@nfv.gwdg.de

***Abstract:** The data management system ECO of the NFV<sup>1</sup> is outlined in terms of matter and technic. The forest site mapping exemplifies the basic principles of the management of soil data, namely integration and standardization of information. Finally, problems that have arisen in operation are addressed.*

***Zusammenfassung:** Das Datenverwaltungssystem ECO der NFV<sup>1</sup> wird allgemein hinsichtlich seiner inhaltlichen und technischen Aspekte skizziert. Am Beispiel der forstlichen Standortskartierung werden Daten-Integration und -Standardisierung als Grundprinzipien der Verwaltung bodenmorphologischer und bodenchemischer Informationen vorgestellt. Abschließend werden im praktischen Betrieb aufgetretene Problemfelder umrissen.*

## **1 Was ist ECO inhaltlich ?**

### **1.1 Messdatenerfassungs- und -verwaltungswerkzeug**

ECO ist ein anwenderorientiertes Datenbanksystem, mit dem vor allem Daten aus ökologischen Erhebungen (Inventuren/Kartierungen, Monitoringprogramme, Versuchswesen) einheitlich erfasst, verwaltet und direkt mit Informationen zur Methodik verbunden werden können. Bei den Messungen kann es sich sowohl um Felderhebungen als auch um Laboranalysen handeln, die ggf. direkt in Beziehung gesetzt werden.

Die zeitliche und räumliche Auflösung sowie die inhaltliche Bedeutung der Messdaten wird datenbankseitig nicht vorgegeben, daher sind diese Kriterien ebenso wie der Detaillierungsgrad der begleitenden Methodendokumentation durch den Benutzer je nach den spezifischen Erfordernissen konfigurierbar.

Durch die Möglichkeit der Fortschreibung sich im zeitlichen Verlauf ändernder Erhebungs- oder Analysemethoden ist ECO besonders für die fortlaufende Speicherung von Messzeitreihen aus längerfristigen Monitoringprojekten geeignet. Aber auch die Verwaltung von episodischen Erhebungen (Inventurdaten mit mehrjährigen Intervallen) ist möglich. Diese werden in Bezug auf den Themenschwerpunkt 'Bodeninformationssysteme' des Workshops unter 5 exemplarisch vorgestellt.

ECO dokumentiert und integriert somit verschiedenste Informationen und die Entstehungsbedingungen von Messdaten. Es ist konzipiert als Datenbasis für externe Auswertungsumgebungen (vgl. 4).

### **1.2 Katalog einschlägiger Geräte und Methoden**

Neben der reinen Messdatenverwaltung kann ECO auch als Verwaltungswerkzeug für Randinformationen genutzt werden, insbesondere zur Katalogisierung von bei der Messdatengewinnung

---

<sup>1</sup> Niedersächsische Forstliche Versuchsanstalt (Forest Research Station of Lower Saxony)

eingesetzten Geräten und Methoden. ECO gibt für diese Bereiche einerseits keine bestimmten Inhalte vor und grenzt andererseits die inhaltlichen Möglichkeiten auch nicht ein. Dies wird ermöglicht über Klassenbildungen (vgl. 2.2), die in identischer Weise auch für andere inhaltliche Bereiche implementiert sind (Geräteinstallationen, Methodenanwendungen, raumbeschreibende Information, Ereignisdokumentation).

### **1.3 Digitales Arbeitsmanual und Projektbuch**

Weiterhin kann ECO zur Normierung von einheitlich zu verwendenden Basisinformationen verwendet werden, wie sie z.B. aus einschlägigen, umweltrelevanten Arbeitsanleitungen (z.B. bodenkundliche Kartieranleitung, Bodenzustandserhebung, forstliche Standortaufnahme etc.) bekannt sind. Darauf aufbauend umfasst ECO die Funktion eines digitalen Projektbuches, in dem solche Basisinformationen zur Beschreibung von Versuchsflächen oder Messstationen verwendet werden können. Hierbei ist die Dokumentation der Veränderungen im zeitlichen Verlauf angesichts der meist sehr langfristigen Monitoringprojekte ein wesentlicher Aspekt.

### **1.4 Steuerungs- und Kontrollinstrument**

Ein grundlegendes Merkmal von ECO besteht in der Bereitstellung spezifischer, projektorientierter Anwendungsprogramme, die die Arbeitsschritte in ihren logischen Abhängigkeiten nachzeichnen und die entsprechenden Funktionalitäten nutzerprofilbezogen anbieten (SCHULZE U. HOPPE 1997). Dadurch wird erreicht, dass Benutzer ohne Datenstruktur- und ohne SQL-Kenntnisse autark mit dem System arbeiten können und dabei einheitliche Konsistenz- und Plausibilitätskriterien zum Einsatz kommen.

## **2 Was ist ECO technisch ?**

### **2.1 Architektur**

Technisch besteht ECO aus drei Ebenen:

- eine abstrakte, kompakte, strukturell stabile Datenbankstruktur liegt auf einem Datenbankserver (50 Tabellen, max. 9 Attribute/Tabelle) und bildet die Basis (HOPPE U. SCHULZE 1997)
- an spezifischen Fragestellungen orientierte Anwendungsprogramme, mit denen die Inhalte der Datenbank ohne Kenntnisse der Datenbankstruktur oder einer Abfragesprache von verschiedenen Arbeitsplätzen im Parallelbetrieb bearbeitet werden können, liegen auf PC-Clients
- 'Virtuelle Tabellen' beschreiben die für die Anwendungen benötigten komplexeren Datenobjekte in Form von Views und Stored Procedures für lesende und schreibende Zugriffe auf der Datenbankstruktur (HOPPE U. SCHULZE 1998)

Die Produktionsdatenbank läuft unter Oracle 8.1.7 EE auf einer Sun Enterprise 450 unter Solaris 2.7. und belegt ca. 7 GB. Es gibt im Rahmen von Machbarkeitsstudien einige Implementierungen auf PC unter NT4.0 und Windows2000 mit Oracle 8.1.6. Die Applikationen laufen unter Windows auf PC und sind mit Borland Delphi 5 entwickelt.

ECO ist eine Eigenentwicklung der NFV. Die Konzeption begann 1993, der Betrieb im Sommer 1996. Bis 1998 standen für Betreuung und Entwicklung 2 Stellen, seitdem noch eine zur Verfügung.

## 2.2 Klassenkonzept des Datenmodells

Hinsichtlich der Gewährleistung einer weitgehenden inhaltlichen Offenheit bei gleichzeitiger struktureller Stabilität wird im Datenmodell weitgehend von den realen Entitäten abstrahiert. Dabei wird ein quasi-objektrelationaler Ansatz verwendet:

- a) Dinge der realen Welt werden zunächst zu Typen geordnet (Klassenbildung). Das Kriterium der Klassenbildung bzw. Abgrenzung unterschiedlicher Typen ist dabei die Definition unterschiedlicher Mengen typenbeschreibender Parameter (vgl. b).
- b) Einer abgegrenzten Klasse (Typ) werden bestimmte beschreibende Parameter zugeordnet, die für alle einzelnen Objekte relevant sind.
- c) Für konkrete Einzelobjekte können dann die Werte der klassenbezogenen Parameter erfasst werden.

## 3 Welche Inhalte verwaltet ECO ?

(Stand Feb. 2002)

- automatisch erfasste Messungen (z.B. Data Logger) von Meteorologie und Bodenhydrologie (ca. 80 Mio. Werte in 15 Minuten-Intervall in z.T. mehr als 10jährigen Zeitreihen an insgesamt ca. 1000 Messorten)
- manuelle Messungen und Beprobungen von Stoffeintrag, Bodensickerwasser, Oberflächenwasser, Streufall, Schnee, Strahlung im Bestand, Luftgaskonzentration (ca. 550.000 Werte in z.T. mehr als 25jährigen Zeitreihen an insgesamt ca. 4800 Messorten)
- chemische Laboranalysen  
Wasser-, Boden-, Pflanzen-, Humusproben  
(ca. 50.000 Proben mit 800.000 Elementergebnissen unter Führung der Raum-Zeit-Zuordnung und Labormethodik in z.T. mehr als 25jährigen Zeitreihen)
- beschreibende Information  
topographische, vegetationskundliche und administrative Beschreibung von standortkundlichen Aufnahmestellen mit Profil- und Horizontbeschreibungen  
(insgesamt ca. 200.000 Angaben für über 25.000 räumliche Objekte zu ca. 1.200 Profilaufnahmen, davon 400 mit chemischer Analytik)
- Arbeitsanleitungen, Geräte und Methoden  
mehrere einschlägige Arbeitsanleitungen mit insgesamt ca. 7.000 Einzelangaben zu ca. 150 Datenfeldern  
ca. 100 Geräte und ca. 650 Labormethoden mit insgesamt ca. 15.000 Einzelangaben
- die beschriebenen Inhalte entstammen folgenden Projekten:
  - Level II-Dauerbeobachtung (UN/ECE ICP Forests) Niedersachsen
  - Forstliche Standortkartierung Niedersachsen
  - Bodenzustandserhebung Niedersachsen (prototypisch)
  - diverse Fallstudien und Versuche:
    - Eichensterben
    - Erstaufforstung
    - Ausbringung von Holzasche
    - Strahlung im Bestand
    - Kalkung Harz
    - UBA-Deposition
    - Waldlandschaft Solling (Forschungszentrum Waldökosysteme, Uni Göttingen)

## 4 Was ist ECO nicht ?

ECO versteht sich wie beschrieben als Dokumentationsdatenbank im Sinne der Schaffung einer verlässlichen, einheitlich strukturierten und normierten Ausgangsdatenbasis, in der nach allgemeinem Verständnis 'unverrechnete' Rohdaten verwaltet werden. Es bietet daher in der Regel (der unter 5 vorgestellte Bereich stellt eher eine Ausnahme dar!) nicht die Funktionalitäten eines Auswertungswerkzeuges, das für bestimmte Datenbereiche spezifische Verarbeitungsschritte der Datenverrechnung oder -aggregation bereitstellt, sondern fungiert als Datenquelle für solche Umgebungen (vgl. aber auch 6).

Auch wenn Laborergebnisse mit ECO verwaltet und Laborproben in ihren zeitlich-räumlich-inhaltlichen Bezug geführt werden können, ist ECO kein Labor-Informationssystem (LIMS) im engeren Sinn, in dem auch Aspekte der Labororganisation oder der labor-internen Qualitätssicherung aufgegriffen werden. Vielmehr wird das LIMS von ECO aus als 'black box' betrachtet, in die einerseits Analysenaufträge übermittelt und aus der andererseits verifizierte Endergebnisse übernommen werden.

## 5 Detailbetrachtung: Verwaltung von Bodeninformation

Seit 1999 wurde die Datenbasis um bodenmorphologische, bodenphysikalische und bodenchemische Daten aus Profilbeschreibungen und -beprobungen erweitert. Dies geschah ohne Modifikation der Datenbankstruktur allein über die Definition zusätzlicher Virtueller Tabellen. Das zugehörige Anwendungsprogramm integriert dabei eine Vielzahl von Funktionen, die für Monitoringprojekte mit Rücksicht auf die Arbeitsabläufe in getrennten Applikationen realisiert sind. Da die Kernfunktion die Erfassung von Profilbeschreibungen und Randinformationen ist, ist die Applikation (bislang) an der Darstellung von Einzelprofilen ausgerichtet (SCHULZE ET AL. 2000). Mittlerweile werden ca. 1.200 standortkundliche Aufnahmestellen geführt, zu denen verschiedene inhaltliche Informationsebenen verwaltet werden.

### 5.1 Informationsebenen

Zu folgenden inhaltlichen Kategorien werden Informationen nach Maßgabe der niedersächsischen forstlichen Standortkartierung verwaltet:

- die topographische und vegetationsgesellschaftliche Beschreibung der Aufnahmestelle selbst sowie die Zuordnung von ökologischer Standortsziffer und Waldentwicklungstypen
- eine Profilbeschreibung mit Charakterisierung des Bodentyps, der Humusform und des Wasserhaushalts
- die stratigraphische und substratbezogene Beschreibung der geologischen Schichtung des Profils
- die einzelnen morphologischen Horizontbeschreibungen (vgl. Abbildung 1)
- die Charakterisierung des Baumbestandes auf Baumarten- und Einzelbaum-Ebene
- die Beschreibung der Bodenvegetation auf Vegetationsschicht- und Arten-Ebene hinsichtlich Artenzahlen und Deckungsgraden
- die Darstellung der Analyseergebnisse aller Proben eines Bodenprofils (Grundwasser, Humus- bzw. Streuauflagehorizonte, Mineralbodenhorizonte bzw. -tiefenstufen) einschließlich der Zuordnung ihrer räumlichen Herkunft (Einzelhorizont, mehrere Horizonte, Tiefenstufe)
- die Dokumentation der angewandten Probenahme- und Analysemethodik

## 5.2 Prinzipien der Datenverwaltung

### 5.2.1 Integration der Teilinformationen

Alle anfallenden Informationen, also die Charakterisierung der Aufnahmeestelle selbst (Topographie, Bestand, Bodenvegetation), die Profilbeschreibung (Geologie und Substratverhältnisse, Horizontfolge, Wasserhaushalt) und die chemischen Analyseergebnisse werden zusammen mit den Methodeninformationen in einer einzigen Datenstruktur gespeichert und liegen damit direkt verknüpft vor. Insbesondere gibt es keine strukturelle Trennung der morphologischen Beschreibungen von den chemischen oder physikalischen Analyseergebnissen, was traditionell meist der Fall ist. Dies setzt die Abbildbarkeit von teils komplexen Zuordnungssituationen voraus (Mischproben aus mehreren Horizonten und/oder mehrere Proben pro Horizont). Die Methodendokumentation scheint vor allem im analytischen Bereich vorrangig, ist aber auch im Bereich Probenahme wichtig und erfahrungsgemäß einer starken zeitlichen Dynamik unterworfen (KÖNIG U. SCHULZE 2002).

Diese Art der Datenverwaltung ermöglicht insgesamt eine zentrale Pflege der Daten durch Vermeidung von Redundanzen, gewährleistet ihre Konsistenz und Interpretierbarkeit, erhöht die Datenqualität und erleichtert Auswertungen.

### 5.2.2 Vollständige Erfassung aller Ausgangsinformationen

Die strukturierte, lückenlose Erfassung aller anfallenden Einzeldaten (Primärdaten) hat vor dem Hintergrund teilweise noch offener Auswertungsstrategien Priorität, denn unterschiedliche Fragestellungen machen i.d.R. einen Rückgriff auf unterschiedliche, nicht aggregierte Ausgangsdaten erforderlich. Unter diesem Aspekt wurden auch einzelne Datenfelder, die traditionell als Einheit aufgefasst werden (z.B. Standortsziffer, Horizontkennung), in Teilinformationen untergliedert, die getrennt verwaltet werden. Die Dokumentation umfasst neben den Daten im engeren Sinne auch die verwendeten Felderhebungs- und Analysemethoden.

Forstamt: FoAWinnefeld		Revier: RfBrüggefeld		Abteilung: 77-m1#43		Aufnahmestelle: 323267		Projekt: KartierungOEK1.1			
Aufnahmestelle (DB)		Profil und Geologie (DB)		Bodenhorizonte		Baumbestand		Vegetationsaufnahme		chemische Analytik und Auswertungen	
Zustand	Auflagehorizonte	Labornummer	Tiefe[cm]	geo./antr.	Neben	geo./antr.	Haupt	Humus-lagerung	AuflageHoriz-durchwurzelginter	AuflageHoriz-übergang	Humus-brechbarkeit
DB	PAMA01_+70_L	P35587/00	+7.0				L	locker_L	keineDurchwurzelginter	scharf	
DB	PAMA01_+60_Of	P35588/00	+6.0				Of	sperrig_Of_Nadel	niedrig	scharf	
DB	PAMA01_+30_Oh	P35588/00	+3.0				Oh	kompakt_Oh	niedrig	scharf	unscharf

Zustand	Mineralbodenhorizonte	Labornummer	Tiefe[cm]	geol.	geo./antr.	Neben	geo./antr.	Haupt	Fein/Sand	Skelett	Lagerung	Silikat-gehalt	Humus-gehalt	Kalk-gehalt
DB	PAMA01_-5_Aeh	B41572/00	-5.0	I				Aeh	Uls	0	Ld1.5			
DB	PAMA01_-35_Ai	B41573/00	-35.0	I				Ai	Uls	0	Ld1.5		h0	
DB	PAMA01_-85_Bvt	B41574/00	-85.0	II		Swd		Bvt	Lu	0	Ld2		h2	
DB	PAMA01_-180_Bbt	B41575/00	-180.0	II		Swd		Bbt	Lu	2	Ld3		h3	
DB	PAMA01_-200_Bv	B41576/00	-200.0	II,III				Bv	Lu	5	Ld3		h4	
DB	PAMA01_->-210_Cv		-210.0	III			ij	Cv	Si4	10	Ld3.5		h5	
									silikatisch, Lockersubstrat				h6	

Abbildung 1: Hauptformular mit horizont-bezogenen Informationen (Ausschnitt)

### 5.2.3 Vorgabe von 'lists of values' für möglichst viele Datenfelder

Die integrierende Datenstruktur ermöglicht die Schaffung eines zentralen Stammdatenpools, auf den alle potentiellen Anwendungen einheitlich zugreifen können. Wo immer möglich werden dabei die Inhalte von Stammdatenfeldern normiert, d.h. mit Listen zulässiger Einträge hinterlegt (vgl. Abbildung 2, links). Freie Benutzereingaben beschränken sich auf die Inhalte nicht formalisierbarer Datenfelder wie z.B. 'besondere Bemerkungen'. Auch dies geschieht unter dem Aspekt einer verbesserten Auswertbarkeit. Denn eine zuverlässige, d.h. komplette Ergebnismenge ist nur lieferbar, wenn ein konkreter Dateninhalt datenbankweit in eindeutiger Schreibweise vorliegt. Dies ist erfahrungsgemäß nur zu gewährleisten, wenn der Begriff vom System selbst vorgegeben wird.

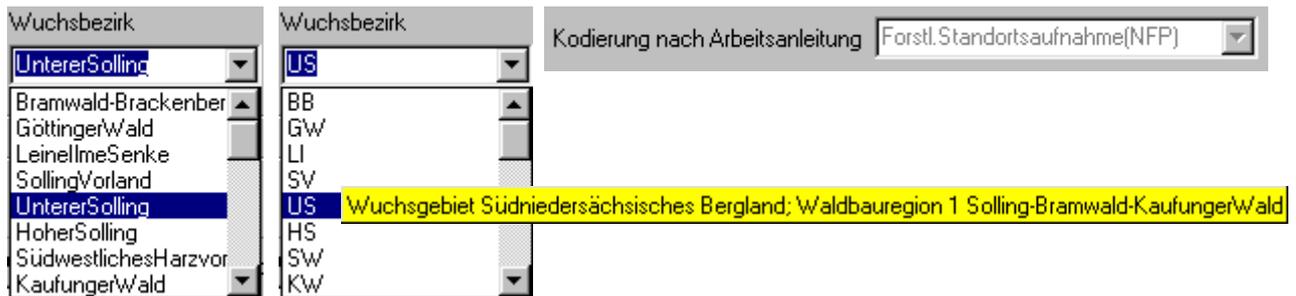


Abbildung 2: Beispiel der Auswahlliste zu einem Datenfeld, das entsprechend einer Arbeitsanleitung kodiert ist und über weitere Erläuterungen der Einzeleinträge verfügt

### 5.2.4 Spezifische Kodierung der Datenfeld-Listen gemäß relevanter Arbeitsanleitungen

Die mögliche Nutzung von Daten durch fachfremde Personen erfordert eine selbsterklärende Wiedergabe von projektintern gebräuchlichen Verschlüsselungen (vgl. Abbildung 2, rechts). Weiterhin muss dem Umstand Rechnung getragen werden, dass es Arbeitsanleitungen gibt, die sich inhaltlich zu einem großen Teil überlappen. Nicht selten werden dabei identische Inhalte in diesen Arbeitsanleitungen verschieden kodiert (vgl. Abbildung 3). Vor diesem Hintergrund wurde in ECO ein Konzept der sog. 'Kodierschemata' umgesetzt, mit dem Parameterlisten einheitlich verwaltet und spezifische Teilmengen spezifisch kodiert auf Projekte zugewiesen werden können.

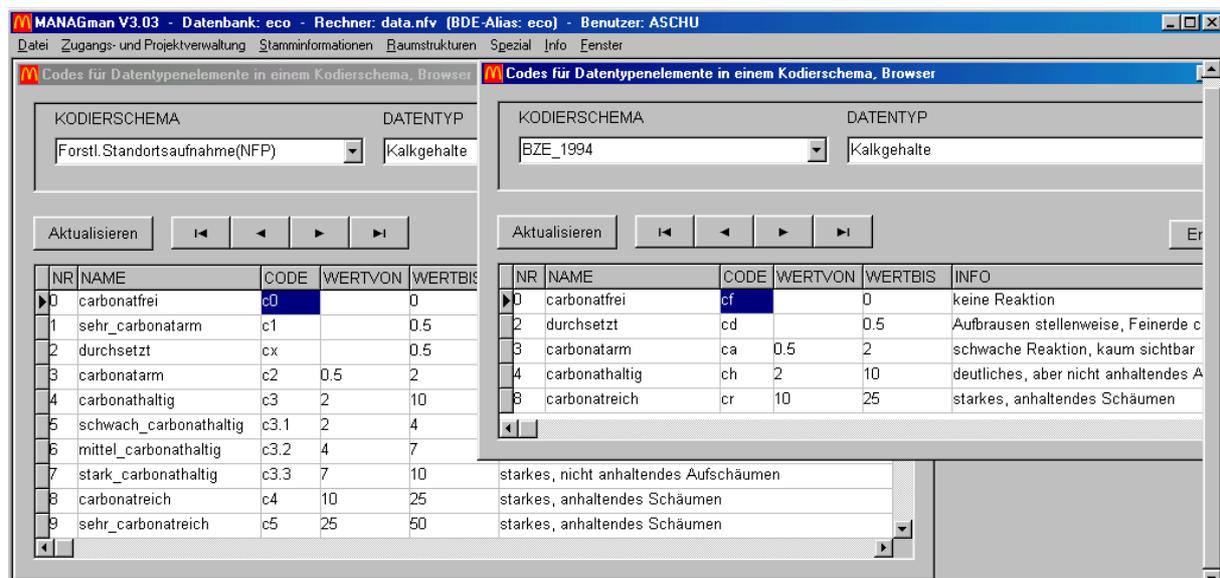


Abbildung 3: Identische Inhalte werden in verschiedenen Arbeitsanleitungen unterschiedlich kodiert. Außerdem werden u.U. nur Teilmengen einer Basisliste verwendet.

### 5.3 Möglichkeiten des Datenzugriffs und der Auswertung

Standortkundliche Aufnahmestellen sind datenstrukturell in die Hierarchie der räumlichen Forstbetriebseinheiten eingebunden (Forstamt - Revier - Abteilung) und können im existierenden Anwendungsprogramm, das die Datenerfassung in den Vordergrund stellt, in dieser Form einzeln ausgewählt, bearbeitet und standardisiert ausgewertet werden (Abbildung 4). Dabei erfolgt i.d.R. auch eine Einstufung in bestimmte Bewertungsrahmen.

- Ableitung bodenphysikalischer Kenngrößen wie Porenvolumen und Trockenraumdichte (falls nicht analytisch erhoben)
- Berechnung der effektiven oder totalen Kationenaustauschkapazität aus Analyseergebnissen
- Zuordnung der pH-Werte zu Pufferbereichen
- Berechnung der Basensättigung
- Berechnung von Elementverhältnissen der Hauptnährelemente
- Berechnung der Elementvorräte für die Auflagehorizonte und das Gesamtprofil

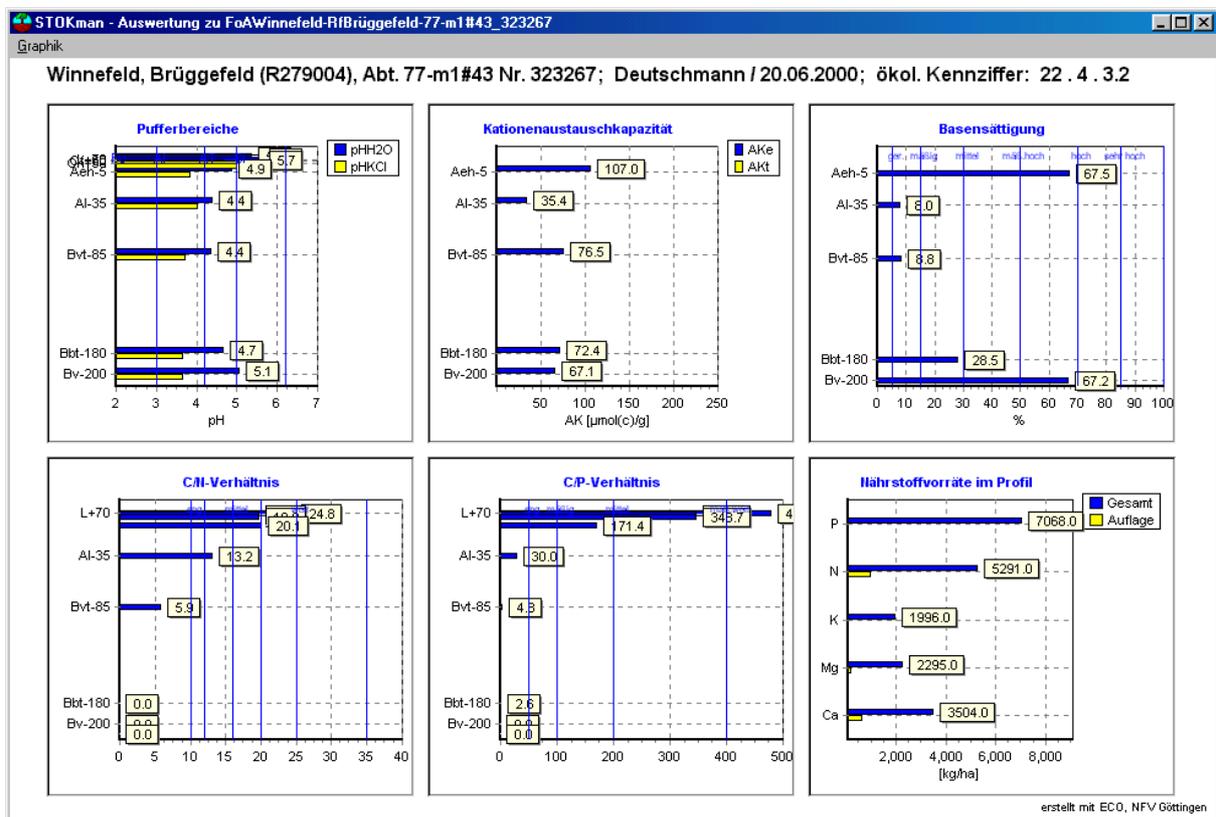


Abbildung 4: Profilbezogene Darstellung von errechneten bodenchemischen Kenngrößen

Prototypisch wurde eine profilübergreifende Abfrageumgebung begonnen. Hier werden beliebige inhaltliche Abfragekriterien formulierbar und kombinierbar, etwa in der Form "suche alle Profile, die im Wuchsbezirk 'xxx' liegen und den Bodentyp 'yyy' haben" (vgl. Abbildung 5).

Eine rein räumliche Identifizierung über Gauß-Krüger-Koordinaten ist ebenfalls vorgesehen, setzt aber die exakte Erhebung der Koordinaten im Feld voraus (z.B. mit GPS). Für eine in dieser Form erweiterte Datenbasis wird die Einbindung einer digitalen kartographischen Oberfläche zur besseren Visualisierung der Raumbezüge und zur einfacheren Navigation möglich, wobei die räumlichen Kriterien mit den inhaltlichen und administrativen verschnitten werden könnten.



## 6 Ausblick

Im Routinebetrieb von ECO an der NFV werden sukzessive bisher noch nicht von ECO abgedeckte Datenbereiche des forstlichen Umweltmonitorings ergänzt. Aktuell ist die Erfassung des Kronenzustands (jährliche Waldzustandserhebung) in Arbeit. Für die Zukunft geplant sind die Integration von Waldernährung und -zuwachs. Parallel soll der hier vorgestellte Bereich der Bodeninformationen gemäß der Erfordernisse unterschiedlicher Projekte universell zur Verfügung gestellt werden.

Im Rahmen des laufenden BMBF-Verbundprojektes "Konzept und Machbarkeitsstudie für die integrierende Auswertung von Daten des Umweltmonitorings im Wald" werden im von der NFV geleiteten Teilprojekt 2 "Datenmanagement" u.a. die unter 5.4 genannten Aspekte im Rahmen einer Überarbeitung des Datenmodells von ECO analysiert und aufgegriffen.

Weiterhin ist eine Erweiterung der Datenbankstruktur hinsichtlich der Verwaltung von Auswertungsdaten sowie der Dokumentation der Ableitungswege (Eingangsdaten, Ersatzdaten, Parameter, Formeln und Algorithmen) in Arbeit.

Es laufen derzeit Machbarkeitsstudien bzgl. der Einsetzbarkeit von ECO im Rahmen des Intensiv-Monitorings bei drei weiteren Bundesländern und der Bundesforschungsanstalt für Forst- und Holzwirtschaft (BFH, Eberswalde) als deutschem Datenzentrum.

## 7 Literatur

HOPPE, J.; SCHULZE, A. (1997). ECO - Datenbank zur Stoffbilanzierung in Waldökosystemen. Teil 1: Datenbankstruktur und Virtuelle Tabellen. *Berichte aus dem Forschungszentrum Waldökosysteme Göttingen, Reihe B 54*, 124 S.

Fortschreibung: <http://www.nfv.gwdg.de/d/eco/eco.html>

HOPPE, J. ; SCHULZE, A. (1998). Virtuelle Tabellen als Werkzeug zur spezifischen Abbildung inhomogener Datenbestände in einem homogenen Datenmodell. In: Hoppe, J. et al. (Hrsg.). *Vernetzte Umweltinformation*. Praxis der Umwelt-Informatik 7, 125-140.

KÖNIG, N.; SCHULZE, A. (2002). Qualitätssicherung im forstlichen Umweltmonitoring durch Methodendokumentation und deren datentechnische Integration am Beispiel der chemischen Analytik. *UBA-Texte*, im Druck.

SCHULZE, A.; HOPPE, J. (1997). Qualitätssicherung bei der Bilanzierung von Stoff- und Wasserflüssen in Waldökosystemen durch datenbankgestützte Arbeitsorganisation. In: Geiger, W. et al. (Hrsg.). *Umweltinformatik '97. 11. Int. Symp. Umweltinformatik Strasbourg 1997*. Umwelt-Informatik aktuell 15, 216-224.

SCHULZE, A.; STÜBER, V.; MEIWES, K. J.; HÖVELMANN, T.; WACHTER, H. (2000). Datenbanklösung zur Verwaltung von Profildaten der forstlichen Standortskartierung. *Forst u. Holz* 55, 110-113.