

# Qualitätssicherung bei der Datenerfassung im FORST-GIS

GEORG LOTHER, MÜNCHEN

## Abstract

*The Geographical Information System of the Bavarian Forest Administration (FORST-GIS) has been operating for seven years now. Meanwhile FORST-GIS has become a large database for forest applications. The establishment of the spatial database requires large investments. The multiple use of the data is mainly determined by its quality, which has to be guaranteed by data capture and data processing. To reach a high level of quality during the building up of the spatial database a determine quality model is necessary. This article describes some methods for quality assurance taking the example of the FORST-GIS.*

## 1 Einführung

Der aktuellen, umfassenden Information über geo-räumliche Strukturen und Zusammenhänge wird in Politik, Wirtschaft und Verwaltung zunehmende Bedeutung beigemessen. Moderne Instrumente zur Verarbeitung raumbezogener (spatial) Daten sind Geoinformationssysteme (GIS). Den Kern eines GIS bilden die GIS-Daten; ihr Informationsgehalt ergibt sich aus der umfassenden Verfügbarkeit für einen Themenbereich und ihrer Qualität. Das Einsatzspektrum der GIS-Daten („Mehrfachnutzung“) wird dabei wesentlich von ihrer Qualität bestimmt. Zieht man in Betracht, dass die Kosten für den Aufbau einer GIS-Datenbank bis zu 80% der Gesamtkosten eines GIS ausmachen können, sollte bei der Konzeption und dem Einsatz eines GIS der Qualitätssicherung für die GIS-Daten ausreichende Beachtung geschenkt werden. Im folgenden werden Möglichkeiten zur Qualitätssicherung von GIS-Daten am Beispiel des Geoinformationssystems der Bayerischen Staatsforstverwaltung (FORST-GIS) aufgezeigt.

## 2 Qualitätsmanagement

Die Strategie zum Erreichen der geforderten Qualität von Produkten hat sich in den vergangenen Jahren von der Qualitätskontrolle hin zum Qualitätsmanagement gewandelt. Während bei der Qualitätskontrolle (Endkontrolle) lediglich fehlerhafte Teile aus der Produktion herausgefiltert, also bereits entstandene Fehler beseitigt werden, ist es das Ziel des Qualitätsmanagements Fehler zu vermeiden, d.h. erst gar nicht entstehen zu lassen. Mit der zunehmenden Komplexität von Produkten, ihrer Entwicklung und ihren Herstellungsprozessen war die oft gebräuchliche Devise: „Wir haben nie Zeit, es gleich richtig zu machen, aber wir haben immer Zeit, es noch einmal zu machen“ nicht mehr durchzuhalten. Die hohen Kosten der Fehlerbeseitigung, aber auch mögliche Image- und Marktverluste für ein Unternehmen bei der Auslieferung fehlerhafter Produkte haben dem Qualitätsmanagement einen rasanten Durchbruch verschafft. Die nötige Basis für den Aufbau von Qualitätsmanagementsystemen wurde in den Jahren 1990/91 mit der Einführung der Normenreihe DIN EN ISO 9000 geschaffen. Diese Normenreihe gibt Rahmenmodelle vor, nach denen für eine Organisation ein geeignetes Qualitätsmanagementsystem eingeführt werden kann, das dieser ermöglicht, die geforderte Qualität von Produkten oder Dienstleistungen zu erreichen und aufrechtzuerhalten.

Unter Qualitätsmanagement versteht man die Gesamtheit aller Maßnahmen in einem Unternehmen, die zur Sicherung der Produktqualität erfolgen. Das Qualitätsmanagement ist eine Aufgabe der Unternehmensleitung. Der dafür erforderliche Apparat wird als Qualitätsmanagementsystem (QMS) bezeichnet. Ein QMS betrifft die gesamte Organisationsstruktur, insbesondere die Verantwortlichkeiten sowie die Verfahren und Methoden zur Qualitätssicherung. Die operativen Regeln eines QMS können mit „Plane, was du tust, tue, was du geplant hast und

schreibe auf, wie du es getan hast“ kurz umrissen werden. Ein wirkungsvolles QMS muss die besonderen Gegebenheiten in einem Betrieb berücksichtigen. Dazu müssen die in der Normenreihe DIN EN ISO 9000 vorgegebenen Rahmenmodelle an die speziellen Bedürfnisse eines Betriebes angepaßt werden. Die Einführung eines QMS ist ein iterativer Vorgang, mit dem Ziel es ständig zu optimieren, um die geforderte Produktqualität bei möglichst hoher Produktivität zu erreichen.

### 3 Qualität im GIS

Ziel des Qualitätsmanagements ist die Sicherung der Qualität von Produkten und Dienstleistungen. Qualität ist gemäß ISO 8402 „*die Gesamtheit von Merkmalen einer Einheit bezüglich ihrer Eignung festgelegte und vorausgesetzte Erfordernisse zu erfüllen.*“ Qualität bezieht sich damit auf objektiv beschreibbare oder meßbare Eigenschaften eines Produktes oder einer Dienstleistung, sie ist also immer im Kontext einer bestimmten Sache zu sehen. Qualität bedeutet dabei nicht, die genaueste und perfektteste Lösung zu erzielen, sondern die geforderten Merkmale einer Einheit so genau wie nötig zu erfüllen.

Für die Beschreibung der Qualität im GIS ist zunächst anzugeben, welche Produkte mit einem GIS hergestellt werden. GIS-Produkte können zusammenfassend als Report bezeichnet werden, wobei ein Report das Ergebnis einer Auswertung ist und sich aus Karten, Tabellen und Texten zusammensetzt. Da mit einem GIS raumbezogene Daten verarbeitet werden, kommt der Karte als Präsentationsmittel für Auswertungen eine vorrangige Bedeutung zu, häufig ist das Ergebnis einer Auswertung ausschließlich eine Karte. Die Qualität eines GIS-Reports hängt dabei von methodischen (Datenverfügbarkeit, Datenqualität, Auswertelgorithmen, Präsentationsmethoden) und organisatorischen Faktoren (Organisation und Qualifikation des Personals, verfügbare Mittel und Geräte) ab

Aus methodischer Sicht ist die Reportqualität von den verfügbaren Daten und den verwendeten Auswerte- und Präsentationsmethoden abhängig. Die Auswertemethoden sind je nach Anforderung sehr vielfältig, sie können von einfachen Berechnungsverfahren bis hin zu komplexen Simulationen reichen. Die Qualität der eingesetzten Algorithmen in Verbindung mit der Qualifikation des Auswerters sind wichtig für ein gutes Ergebnis. Die Daten haben jedoch die Schlüsselstellung für die Reportqualität, denn mit veralteten oder schlechten Daten kann sie auch durch noch so gute Auswerte- und Präsentationsmethoden nicht wesentlich verbessert werden.

Bei GIS-Daten ergibt sich die Datenqualität aus der im Datenmodell definierten nominellen Wirklichkeit und aus der Sorgfalt bei der Datenerfassung. Im Datenmodell müssen die Objektklassen und ihre Beziehungstypen so beschrieben sein, dass die Objekte über alle nötigen Merkmale verfügen. Versäumnisse bei der Datenmodellierung können später nur mit großem Aufwand ausgeglichen werden. Hier soll jedoch nur auf die Qualitätssicherung bei der Datenerfassung eingegangen werden. Dazu ist es erforderlich für GIS-Daten objektive Kriterien - Qualitätsmerkmale - anzugeben, die eine hinreichende Beurteilung ihrer Qualität zulassen. So kann z.B. für ein Flurstück die Qualität seiner Flächenangabe (Merkmal des Objektes Flurstück) durch die Angabe der Standardabweichung (Qualitätsmerkmal) beschrieben werden. Mittels Vergleich der Standardabweichung mit der vorgegebenen Fehlergrenze läßt sich dann prüfen, ob die geforderte Qualität erreicht wurde.

Für die Modellierung von GIS-Objekten werden zusätzlich zu den in anderen Informationssystemen üblichen semantischen und temporalen Merkmalen geometrische Elemente benötigt. Für diese Merkmalstypen werden in der Literatur (siehe z.B. Caspary 1993) meist fünf bis sechs Eigenschaften angegeben, deren Prüfung und Dokumentation für die Qualitätssicherung von GIS-Daten erforderlich sind. Die Qualitätsmerkmale sollen dem Benutzer von GIS-Daten hinreichende Auskunft darüber geben können, welche Qualität er für die Aussagen eines Re-

ports zu erwarten hat bzw. inwieweit die Daten für seine Aufgaben überhaupt geeignet sind (Ausschlusskriterien). Sie sollen dabei jedes für sich einen möglichst unabhängigen Bereich eines Datenbestandes abdecken und gemeinsam dazu geeignet sein die Qualität aller möglichen GIS-Objekte vollständig zu beschreiben; sie ergeben das Qualitätsmodell für einen GIS-Datenbestand.

### Qualitätsmodell für GIS-Daten

Geltungsbereich	Qualitäts-Merkmal	Beschreibung
Gesamter Datensatz Themenbereich Gebietsfragment	pragmatische Qualitätsmerkmale	Freier Text expliziter Verweis impliziter Verweis
	<b>Herkunft</b> <b>Aktualität</b> <b>Datenbeschaffenheit</b>	
Objekt Objektteil: Geometrie Objektteil: Attribut	deskriptive Qualitätsmerkmale <b>metrische Genauigkeit</b> <b>semantische Richtigkeit</b> <b>logische Konsistenz</b>	Kenngröße Bedingung Qualitätsfolie

Die sechs Qualitätsmerkmale für GIS-Daten kann man in pragmatische und deskriptive Qualitätsmerkmale einteilen.

Die pragmatischen Qualitätsmerkmale geben allgemeine, anwendungsorientierte Hinweise über einen Datenbestand. Sie beinhalten i.d.R. die Ausschlusskriterien für seine Verwendung und lassen sich meist schnell beurteilen. Die pragmatischen Merkmale können mittels freiem Text oder Verweis beschrieben werden, sie gelten entweder für den gesamten Datensatz, einen Themenbereich oder ein Gebietsfragment und beinhalten damit die globalen Aussagen über die Qualität der Daten.

Die deskriptiven Qualitätsmerkmale geben Auskunft über die Qualität einzelner Merkmale, Objekte oder Gruppen von Objekten. Sie müssen durch geeignete Tests bestimmt werden. Hierfür sind Prüfmittel (Indikatoren) und Kenngrößen erforderlich, deren Einsatz oft sehr aufwendig werden kann. Für ein Qualitätsmerkmal ist sein Geltungsbereich anzugeben und eine geeignete Form für seine Beschreibung zu wählen. Es werden Qualitätsmerkmale benötigt, die Aussagen über die metrische Genauigkeit und semantische Richtigkeit von Objektmerkmalen sowie über die logische Konsistenz der Beziehungen von Objektteilen untereinander und von Objekten miteinander machen. Die deskriptiven Merkmale haben meist lokalen Charakter, sie beziehen sich auf Objektteile, Objekte oder Objekttypen und lassen sich mit Verweisen, Kenngrößen oder (topologischen) Bedingungen beschreiben. Bei der Datennutzung sollte die Prüfung der deskriptiven Qualitätsmerkmale - wegen des Aufwands - erst dann erfolgen, wenn die Daten aus pragmatischer Sicht für einen Report überhaupt geeignet sind.

### 4 Daten-Qualität im FORST-GIS

Das FORST-GIS ist ein topographisches Fach-Geoinformationssystem für die Bestandsdokumentation der forstlichen Flächeninformation (Lother und Rottman 1999, Lother 1999). Die pragmatischen Qualitätsmerkmale im FORST-GIS unterscheiden sich im Wesentlichen nicht von denen anderer Datenbestände. Es wird deshalb nur auf die Ermittlung von deskriptiven Qualitätsmerkmalen im FORST-GIS eingegangen. Folgende Prüfmittel zur Ermittlung deskriptiver Qualitätsmerkmale werden eingesetzt:

- **Metrische Qualität:** Überbestimmte Transformation, unabhängige Kontrollpunkte und Kontrollflächen, Vergleich unabhängiger Datenquellen
- **Logische Konsistenz:** topologische Bedingungen in Kombination mit Attributen, Prüfung der referenziellen Integrität

- **Semantische Richtigkeit:** unabhängige, redundante Erfassung von Attributen

Im FORST-GIS wird versucht, die Qualitätsmerkmale mit möglichst geringem Aufwand zu dokumentieren, um einerseits den Datenbestand nicht unnötig aufzublähen und andererseits den zusätzlichen Erfassungsaufwand in vertretbaren Grenzen zu halten. Meist werden implizite Verweise (Zuordnung von Genauigkeiten zu Attributen) verwendet. Eigene Attribute (explizite Verweise) in der Datenbank werden nur dann eingesetzt, wenn sie auch eine weitgehend selbstständige Bedeutung haben (z.B. abgeglichene Fläche eines Bestandes).

Für die Ermittlung der Qualitätsmerkmale werden die kontrollierte Redundanz und die Prüfung von topologischen Bedingungen eingesetzt. Die redundante Information wird meist nur während der Datenerfassung benötigt und muß nicht in die Datenbank eingehen. Ziel ist es einen weitgehend redundanzfreien Datenbestand aufzubauen, dessen Qualität durch Qualitätsmerkmale und dokumentierte Prüfungen nachgewiesen wird. Bei der Erfassung von GIS-Daten, ist insbesondere für Aussagen über ihre metrische Genauigkeit und Attributrichtigkeit, unabhängig erfaßte, redundante Information m.E. unbedingt erforderlich. Eine redundanzfreie Datenerfassung führt zwar direkt zu einem redundanzfreien und damit widerspruchsfreien Datenbestand, dieser ist aber auch unzuverlässig, da er nicht kontrolliert wurde oder werden kann, d.h. Fehler können lediglich nicht aufgedeckt werden, da es keine Kontrollmöglichkeiten gibt - widerspruchsfrei bedeutet nicht fehlerfrei!

## 5 Zusammenfassung

Die Prüfung und Dokumentation der Qualität von GIS-Daten ist für die Abschätzung der Reportqualität Voraussetzung. Die Qualitätsmerkmale sollten geeignet sein, daß sie im Sinne des Fehlerfortpflanzungsgesetzes auf Auswertungen übertragen werden können, um die Genauigkeit und Zuverlässigkeit der Aussagen eines Reports beurteilen zu können. Die pragmatischen Merkmale eines Datenbestandes müssen unbedingt beschrieben sein, bei den deskriptiven Merkmalen sollten Minimalanforderungen erfüllt werden. Qualitätsmerkmale wie z.B. Lagestati von Punkten, Pufferbreiten für Linien, Aussagen über Stichprobenvergleiche mit unabhängigen Kontrollgrößen lassen sich meist mit vertretbarem Aufwand ermitteln und sichern die Qualität der GIS-Daten.

## 6 Literatur

- CASPARY, W. (1993): Qualitätsaspekte bei Geoinformationssystemen. Zeitschrift für Vermessungswesen Jg. 124, Heft 8/9, S. 444-449
- LOTHER G., ROTTMANN M. (1999): Stand, Ziele, Perspektiven und Planungen auf dem Gebiet der Forstkartografie in der Bayerischen Staatsforstverwaltung, Schweizerische Zeitschrift für Forstwesen Jg. 150, H.2, S. 56-66
- LOTHER G. (1999): Geoinformationssysteme - Begriffe und Kategorien, GEO-Informationssysteme Jg. 12, H.6, S.20-27