

SAMT – eine neue Open Source Plattform zur Landschaftsanalyse, Modellentwicklung und Integration räumlich expliziter ökologischer und ökonomischer Modelle

Ralf Wieland; Marion Voss; Wilfried Mirschel; Karl-Otto Wenkel; Igbekele Ajibefun

Zentrum für Agrarlandschafts- und Landnutzungsforschung (ZALF) Müncheberg,
Institut für Landschaftssystemanalyse
Eberswalder Str. 84, D-15374 Müncheberg
rwieland@zalf.de

Abstract: Das *Spatial Analysis and Modeling Tool* (SAMT) ist ein neuartiges gridbasiertes Modellierungsinstrument und wurde als Open Source Projekt am ZALF entwickelt. SAMT besitzt einige Eigenschaften eines GIS in Form grundlegender Gridfunktionen, hat neben der IO-Schnittstelle zum GIS eine eigene Datenverwaltung und ist auf die problemlose Integration unterschiedlicher Modelltypen (dynamische Modelle, neuronale Netze, Fuzzysysteme etc.) optimiert. Vorgestellt werden die Struktur, Eigenschaften von SAMT und sich daraus ergebenden Möglichkeiten der Simulation. Auf praktische Anwendungen wird in einem gesonderten Beitrag eingegangen.

1 Motivation

Die Nachhaltigkeit einer Landschaft hängt bekanntermassen von einer Vielzahl von Faktoren ab. So geht das "**Drei-Säulen-Modell**" des BDI [bdi04] von der Vorstellung aus, dass Nachhaltigkeit durch das gleichzeitige und gleichberechtigte Umsetzen von umweltbezogenen, sozialen und wirtschaftlichen Zielen erreichbar ist. Für die Entwicklung räumlicher Simulationssysteme, die eine Analyse der Nachhaltigkeit einer Landschaft als Zielstellung haben, ergibt sich die Forderung der Integration von Modelltypen, die soziale, ökonomische und Umweltaspekte abbilden. Das erfordert die Integration unterschiedlicher Simulationsverfahren. Um die Interaktionen zwischen Teilflächen der Landschaft abbilden zu können, muss die Simulation räumlich konkret und in den meisten Fällen dynamisch sein. Ein weiteres Merkmal, das im Kontext der Nachhaltigkeit auftritt, ist, dass bestimmte Beziehungen innerhalb des Systems nur ungenau bekannt sind bzw. die Inputs nur unscharf vorliegen. Letzteres impliziert die Verwendung von Fuzzymodellen oder auch neuronalen Netzen als Mittel der Modellierung. Für die Simulation räumlicher Zusammenhänge kommen heute geographische Informationssysteme (GIS), wie ARCGIS [gisa04] oder als freie Alternative GRASS [gisb04] zum Einsatz. Das Hauptziel bei der Entwicklung dieser Systeme liegt in der Speicherung räumlicher Daten, der Durchführung räumlicher Analysen und der Präsentation von Analyseergebnissen in Form von Karten. Es handelt sich dabei um leistungsfähige Softwareprodukte, die ausgezeichnete Ergebnisse produzieren, aber auch ihren Preis haben. Neben dem hohen Preis ist nachteilig, dass sie nur wenig Unterstützung für die Modellintegration

und Simulation bieten und dass ihre Handhabung recht kompliziert ist. So gibt es beispielsweise keine einfache Möglichkeit, Fuzzymodelle oder neuronale Netze in die bestehenden GIS zu integrieren. Auch die Integration komplizierter räumlicher Ableitungen oder gar dynamischer Simulationen wird nur ungenügend unterstützt. Bedingt durch die Interpretation der mit dem GIS gelieferten Skriptsprache ergeben sich Laufzeiten, die für eine interaktive Arbeit nicht akzeptabel sind. Etwas anders stellt sich die Situation bei der Nutzung von Simulationstools, wie Matlab [mat04] oder den freien Alternativen wie Octave [octave04] dar. Hier liegen leistungsfähige Methoden als kompilierter Code vor, die dann durch eine Skriptsprache gesteuert werden. Diese Systeme, die auf die Simulation spezialisiert sind, besitzen ein ausgezeichnetes Laufzeitverhalten, wenn auf die vorgefertigten Methoden zurückgegriffen werden kann. Nachteilig ist hier vor allem die fehlende Schnittstelle zu einem GIS (oder allgemeiner zu geographischen Daten) und das Nichtvorhandensein spezieller räumlicher Analyse- und Simulationstechniken. Zu bemerken ist, dass die Nachbildung räumlicher Simulationen durch die zugehörige Skriptsprache zu ähnlichen Laufzeitproblemen wie bei den einzelnen GIS führt. Aus der Analyse dieses Zustandes entstand die Notwendigkeit, ein Simulationstool zu schaffen, das speziell die räumliche Simulation unterstützt.

2 Allgemeine Charakteristik von SAMT

Zum gegenwärtigen Entwicklungsstand umfasst SAMT folgenden Funktionsumfang:
Allgemeine Charakteristik:

- rasterbasierte Simulation (basierend auf einer eigenentwickelten Grid-Library)
- wichtige Grundalgorithmen zur Rasterverarbeitung (Add, Mul, Norm,...)
- Zusatzgridfunktionen für die Simulation (Moving-Window, Distance, Interpolation)
- Analysefunktionen (Histogramm, 3D-Sichten, Splattertechnik, Hassediagramme,...)
- schnelles HDF-Format für interne Nutzung und Modellkopplung
- Steuerung über ansprechendes graphisches Nutzerinterface.

Die rasterbasierte Simulation hat sich als leistungsstark und für viele Prozesssimulationen, wie die Wasser- und Winderosion, oder die Modellierung von Sukzessionsprozessen als einzig gangbarer Weg erwiesen. Dafür wurde eine Reihe von Grundalgorithmen implementiert. Zu diesen gehören beispielsweise die Möglichkeit der Multiplikation eines Grids mit einer reellen Zahl, die Normierung eines Grids in den Bereich [0..1], die Klassifikation von Grids etc. Zusätzlich wurden die additive und multiplikative Überlagerung zweier Grids, die Berechnung von Distanzen etc. integriert. Als Analysetechniken werden neben z.B. Histogramm und 3D-Ansicht auch weniger bekannte Techniken wie die Splattertechnik [SML98] (mehrdimensionale Datenanalyse) oder die Hassediagrammtechnik [Wi04], die zur Modellvalidierung herangezogen wird, eingesetzt.

Schnittstellen:

- Schnittstelle zum GIS ARC/INFO via ASCII-Grid
- Schnittstelle zum Datenbanksystem MySQL
- Import von Punktdaten im ASCII-Format

- Simulationsrealisierung über externe Prozesse (Kopplung via HDF oder Sockets)
- In- und Export von Bilddaten

Simulation:

- Integration von Clusteralgorithmen und Kohonennetzen als Datenanalysetool (Grundlage für neuronales Netz mit radial basis Functions (RBF))
- Direkte Editierung von Fuzzy Modellen und Anwendung auf räumliche Daten mittels SAMT-integrierter Fuzzy-Toolbox
- Nutzung einer auf den Nutzer bezogenen erweiterbaren Modellbasis bzw. Unterstützung bei Modellierung in C++ durch ein Template.

3 Struktur von SAMT

SAMT besteht aus einem graphischen Nutzerinterface und einer darunter liegenden Gridlibrary, in der die eigentlichen Grid-Operationen ausgeführt werden. Beide Komponenten bilden das Grundgerüst, um das sich eine Reihe weiterer Komponenten gruppieren. Hierzu zählen z.B. die 3D-Analyse, die Fuzzy-Toolbox und die Hassediagrammtechnik. Diese Komponenten bilden feste Bestandteile von SAMT, obwohl sie zum Teil als eigenständige Prozesse eher lose mit SAMT gekoppelt sind. Für den Modellierer wird eine Schnittstelle bereitgestellt, die eine einfache Integration der nutzerspezifischen Modelle ermöglicht. Diese Modelle laufen prinzipiell als eigenständiger Prozess und nutzen als Schnittstelle die Übergabe über das HDF-Filesystem. Dieses Verfahren ist einerseits schnell genug, um die meisten Operationen nicht signifikant zu verlangsamen, andererseits trägt dieses Verfahren zur Stabilität von SAMT bei, da ein Absturz eines Anwenderprogramms nicht zur Beeinträchtigung von SAMT führt. Neben dieser für den Anwender entwickelten Schnittstelle gibt es eine Schnittstelle über Sockets zum Netzwerk. Diese wurde exemplarisch bei der Integration der Hassediagrammtechnik verwendet. Ein weitere Einsatzfall dieser Schnittstelle ist die Auslagerung rechenzeitintensiver Simulationen auf leistungsfähigere Rechner im Netzwerk oder auf ein Cluster. Die Schnittstelle zum freien Datenbanksystem MySQL wird für die Bereitstellung von Sachdaten genutzt. Vorteil dieser serverbasierten Lösung ist die Möglichkeit des Zugriffs durch den Nutzer via ODBC. Als Frontend kommt häufig MS-Access zum Einsatz. Der Systemnutzer kann in seiner gewohnten Umgebung Sachdaten bereitstellen und verändern. Diese Daten stehen dann konsistent der Simulation in SAMT zur Verfügung. So ist gewährleistet, dass stets aktuelle Daten in die Simulation einfließen und dass die Daten für alle Simulationen immer gleich sind. Daneben kann die Datenbankschnittstelle auch zur losen Kopplung von Modellen unterschiedlichen Typs mit den räumlichen Simulationsmodellen eingesetzt werden. In SAMT sind eine Reihe von Methoden integriert, wie z.B. Clusteralgorithmen oder Kohonennetze. Diese Algorithmen sind Bestandteil des Kerns von SAMT und werden schrittweise erweitert. Ziel ist es, auch den Anwender mit in die Erweiterung dieses Teils von SAMT einzubeziehen. So ziehen Einsatzfälle, die eine dreidimensionale Berechnung erfordern (z.B. Winderosion), eine dreidimensionale Visualisierung nach sich. In diesem Fall werden dann entsprechende 3D-Visualisierungssysteme entwickelt. Ein weiterer wichtiger Einsatzfall der 3D-Visualisierung ergibt sich bei der Datenanalyse. Beispielsweise können Punktdaten mittels *Non uniform-*

rational-basisplines (NURBS) [EU96] durch die Krümmung einer Fläche wertvolle Hinweise liefern, wie die Parameter der rechenzeitintensiven Moving-Window-Technik [KL04] einzustellen sind. Insgesamt geht es auch um eine Nutzung moderner Visualisierungstechniken in der Modellierung, d.h. Techniken der Visualisierung werden als Techniken für die Modellierung eingesetzt. Eingang in den Kern von SAMT können allerdings nur Algorithmen erhalten, die ausreichend getestet und die für viele Interessenten nutzbar sind, anderenfalls sind sie nur nutzerbezogene Modelle. Abschließend einige Bemerkungen zur Softwarebasis von SAMT und der angedachten Weiterentwicklung. SAMT wurde als freie Software entwickelt und steht im Internet zum freien Download bereit (<http://www.zalf.de/lsad/persweb/wieland/welt/samt1.html>). Neben den eigenen Softwaremodulen wurde auf eine Reihe bewährter Bibliotheken: C++ application framework QT [qt04], hierarchische Datenformat HDF [hdf04], gnu scientific library [gsl04], Visualisierungssoftware VTK [vtk04] (wird zukünftig ersetzt durch VRS [vrs04]) und Datenbankmanagementsystem MySQL [mysql04] zurückgegriffen. Da SAMT unter Linux lauffähig ist, Linux aber auf dem Desktop noch nicht die große Verbreitung gefunden hat, wurde eine CD entworfen, die zum Testen von SAMT genutzt werden kann. Diese CD basiert auf der Knoppix-CD [knopp04] und kann ohne die Festplatte zu verändern gebootet werden. Da diese CD auch die Quellen von SAMT enthält, kann über den auf der CD enthaltenen Installer Knoppix auf der Festplatte installiert und SAMT im vollen Umfang genutzt werden.

Literaturverzeichnis:

- [bdi04] <http://de.wikipedia.org/wiki/BDI>
- [EU96] Engeln_Müllges G.; Uhlig F.: Numerical Algorithms with C, Springer, Berlin, 1996, 596p
- [gisa04] <http://www.esri.com/software/arcgis/>
- [gisb04] <http://grass.baylor.edu/>
- [gsl04] <http://sources.redhat.com/gsl/>
- [hdf04] <http://hdf.ncsa.uiuc.edu/>
- [KL04] Kiesel, J.; Lutze, G.: Skalierbare Generalisierungsmethoden zur Ausgrenzung von Landschaftseinheiten, IÖR-Schriften 43-2004, Sächsisches Druck- und Verlagshaus AG, 2004, S. 47-63
- [knopp04] <http://http://www.knopper.net/knoppix/>
- [mat04] <http://www.mathworks.com/>
- [mysql04] <http://www.mysql.com>
- [octave04] <http://www.octave.org>
- [qt04] <http://www.trolltech.com/>
- [SML98] Schröder W.; Martin K.; Lorensen B.: "The Visualization Toolkit" Prentice-Hall, London 1998, 645 p.
- [vrs04] <http://www.vrs3d.org/>
- [vtk04] <http://public.kitware.com/VTK/>
- [Wi04] Wieland R.; Brüggemann R.; Simon U.; Voß M.: Kopplung von Ranking Methoden und GIS am Beispiel von Habitatmodellierung, Gnauck (Hrsg): Theorie und Modellierung von Ökosystemen, Shaker Verlag, 2004, (im Druck)