Gesellschaft für Informatik, Arbeitskreis 5
Arbeitskreistreffen 5./6.11.1992
KfK-Bericht (im Druck)
Kernforschungszentrum Karslruhe

Anforderungen an die Dokumentation mathematischer Beschreibungen ökologischer Prozesse

Joachim Benz[1)]
Michael Knorrenschild[2)]

1) Gesamthochschule Kassel · Universität
Fachbereich Landwirtschaft
Abteilung Pflanzenbau II
Nordbahnhofstr. 1a
D-3430 Witzenhausen
e-mail: benz@gsf.de

2) GSF - Forschungszentrum für Umwelt und Gesundheit
Projektgruppe Umweltgefährdungspotentiale von Chemikalien (PUC)
Ingolstädter Landstr. 1
D-8042 Neuherberg
e-mail: knorren@gsf.de

1. Motivation und Ziele

In den vergangenen Jahrzehnten wurde weltweit in erheblichem Umfang versucht, ökologische Prozesse mit Hilfe von mathematischen Modellen zu beschreiben. Die entwickelten Modellansätze stellen eine umfangreiche Ansammlung von Wissen über die Eigenschaften und Strukturen dieser Prozesse dar. Zudem gewinnt die Anwendung der mathematischen Simulation in Planung und Analyse zunehmend an Bedeutung.
Dieses Wissen kann aber nur dann genutzt werden, wenn es

- mit vertretbarem Aufwand verfügbar und
- vollständig dokumentiert ist.

Darüber hinaus ist es wünschenswert, daß unterschiedliche mathematische Beschreibungen ein und desselben Prozesses einem wertenden Vergleich zugänglich sind.
Diese Anforderungen führen auf eine vollständige Dokumentation, die

- die Beschreibung der verwendeten Gleichungen,
- die Beschreibung der Gültigkeitsgrenzen und Randbedingungen sowie
- die Beschreibung der Güte innerhalb der Gültigkeitsgrenzen

umfaßt.
Die gegenwärtige Situation entspricht nur in wenigen Fällen diesen Anforderungen. Die einzelnen Modellansätze unterscheiden sich z.T. bezüglich ihrer Frage- bzw. Aufgabenstellung, in den Zeit- und Raumskalen, dem Detaillierungsgrad sowie ihren Gültigkeitsgrenzen. Diese Unterschiede und/oder Einschränkungen der Aussagefähigkeit sind dem Außenstehenden nur in wenigen Fällen offensichtlich. Vielmehr ist meistens ein genaueres Studium der Modellan-

sätze sowie der entsprechenden Einbettung der Entwicklungsarbeiten notwendig, um sich diese Information zu beschaffen. Darüberhinaus fehlen in der Dokumentation der mathematischen Beschreibungen vielfach Angaben zu den Gültigkeitsgrenzen, den wichtigsten Annahmen und den Eigenschaften bezüglich der Prozessumgebung. Aufgrund der Unterschiede in den Modellansätzen ist die unmittelbare Vergleichbarkeit nur in einzelnen Fällen oder nur eingeschränkt möglich. Die Publikation der Modellansätze bzw. ihrer Anwendungen erfolgt überwiegend in Zeitschriften unterschiedlichster Fachrichtungen.
Dies bedeutet, daß das oben angesprochene Wissen in heterogener Weise, verteilt und zum Teil nur unvollständig vorliegt. Die Folge dieser eingeschränkten Verfügbarkeit ist, daß die Kenntnisse z. Zt. nicht optimal und umfassend genutzt werden können.

Ziel des Forschungsvorhabens ECOBAS ist es, die Verfügbarkeit dieses Wissens zu verbessern. Es soll eine Datenbank aufgebaut werden, in der Dokumentationen mathematischer Beschreibungen ökologischer Prozesse nach bestimmten Gesichtspunkten gespeichert und abrufbar sind.
Um die Datenbankstruktur festlegen zu können, wurde zunächst ein Konzept erarbeitet, das eine vollständige Dokumentation mathematischer Beschreibungen ökologischer Prozesse gewährleistet. Insbesondere wurde auch dem Aspekt Rechnung getragen, dem späteren Benutzer bei der Auswahl für ihn geeigneter Beschreibungen möglichst weitreichend Unterstützung zu bieten.

Im einzelnen wurden folgende Ziele bzw. Ansprüche der Konzepterstellung zugrunde gelegt.

(1) leichte Verfügbarkeit der Information

(2) vollständige und eindeutige Dokumentation der mathematischen Formulierung

(3) möglichst vollständige Dokumentation der Gültigkeitsgrenzen

(4) Berücksichtigung der vernetzten Struktur ökologischer Prozesse
a) Aggregationshierarchie
b) horizontale Vernetzung

(5) Dokumentation durchgeführter Validierungen

(6) Spezifikation der Modellgüte innerhalb des Gültigkeitsbereiches

(7) Dokumentation von Systemeigenschaften innerhalb des Gültigkeitsbereiches

Die zum jetzigen Zeitpunkt vorliegende Konzeption berücksichtigt die Ziele 1-4, die Ziele 5-7 sollen in der weiteren Entwicklungsarbeit integriert werden.

Die Entwicklung des Dokumentationsschemas wird in Kooperation bzw. Abstimmung mit dem Forschungsprojekt 'Umweltforschungsinformationssystem' (UFIS), das vom GSF - Forschungszentrum für Umwelt- und Gesundheit in München bearbeitet wird, und mit dem Projektzentrum Ökosystemforschung der Universität Kiel durchgeführt.

2. Erläuterungen zum Dokumentationsschema

Für einen bestimmten Prozesstyp (z.B Primärproduktion, Photosynthese u.a.) können eine oder mehrere allgemeine mathematische Beschreibungen vorliegen.
Unter der Aggregationshierarchie der Prozesstypen wird hier folgender Sachverhalt verstanden: Viele ökologische Prozesse setzen sich wiederum aus einer Anzahl von Teilprozesse zusammen. Bespielhaft sei hier die Primärproduktion in aquatischen Systemen genannt, die sich aus den Teilprozessen Nettoassimilation, Sedimentation und Mortalität sowie dem 'grazing' zusammensetzt. Als weitere Detaillierung lassen sich z.B. für die Nettoassimilation die Teilprozesse Photosynthese und Respiration anführen. Die Abbildung der Aggregationshierarchie ökologischer Prozesse wird mit Hife des Konzepts von 'embedded processes' realisiert.
Unter horizontaler Vernetzung ist zu verstehen, daß die Ausgangs- bzw. Zustandsgrößen eines Prozesses Eingangsgrößen eines anderen Prozesses sein können.

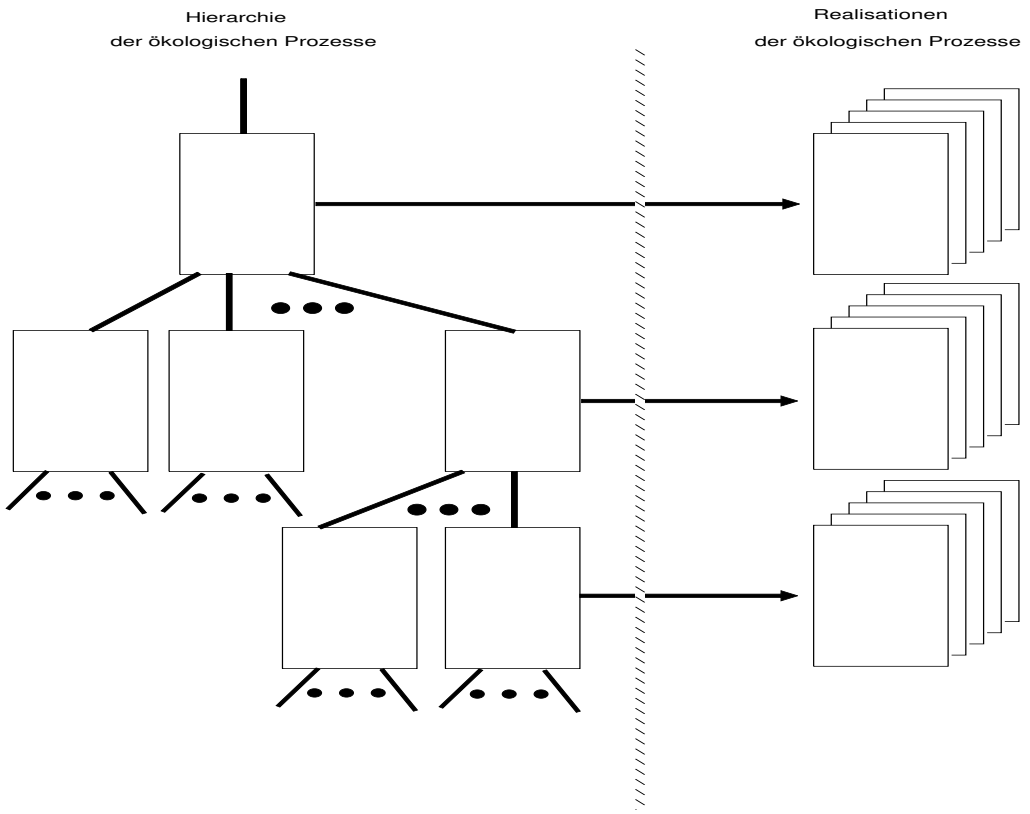

Abb. 1: Struktur der Dokumentation mathematischer Beschreibungen
Für jede allgemeine mathematische Beschreibung eines Prozesstyps existieren eine oder mehrere Realisationen. Hierunter sind die konkreten Verwirklichungen des Prozesstyps für bestimmte Ökosysteme, Organismen, Bodentypen u.a. zu verstehen. Dementsprechend muß die Dokumentation der Realisation die Angabe von Parameterwerten, die Zuordnung zu bestimmten Kategorien (Ökosystemtyp, Klimazone, Organismenspezifikation u.a.), die die Prozessumgebung charakterisieren, enthalten. Ferner sind spezifische Angaben der Umgebungseigenschaften und die zulässigen Wertebereiche des Anfangszustands, der Zustandsgrößen, der Ein- und Ausgangsgrößen, die den Gültigkeitsbereich der Realisation bestimmen, notwendig. Zusätzlich werden hier Informationen bezüglich der Modellentwickler, der zugehörigen Literatur sowie Diskussionsbeiträge gespeichert.
Die sich daraus ergebende übergeordnete Struktur ist in Abbildung 1 dargestellt.

Das vollständige Dokumentationsschema ist im Anhang angegeben.

2.1 Erläuterungen zur Dokumentation eines Prozesstyps

Ein Prozesstyp wird zunächst durch die Bezeichnung, durch eine Liste von Schlagworten und durch eine kurze Beschreibung gekennzeichnet. Die eigentliche Dokumentation des Prozesstyps besteht aus vier Abschnitten. Im Deklarationsteil sind die Eingangsgrößen, die Zustandsgrößen sowie die Parameter anzugeben. Unter Zustandsgrößen werden hier alle Größen, die den Zustand eines Systems beschreiben, verstanden. Die hier benutzte Definition umfaßt sowohl Zustandsgrößen als auch Ausgangsgrößen im Sinne der klassischen Systemtheorie (Zustandsraumdarstellung).
Im 2. Abschnitt (equations) sind die Gleichungen anzugeben. Dieser Abschnitt des Dokumentationsschemas ist so konfiguriert, daß die wichtigsten Funktionsklassen erfaßt werden können. Unterschieden wird dabei zwischen partiellen Differentialgleichungen (PDE's), dynamischen Zustandsgleichungen und algebraischen Gleichungen. Dynamische Zustandsgleichungen werden darüberhinaus nach der Eigenschaft der Zeitskala (kontinuierlich, diskret), der Eigenschaft der Zustandsgröße (kontinuierlich, diskret) und ob es sich um einen deterministischen oder stochastischen Ansatz handelt, unterschieden. Für PDE's und die dynamischen Zustandsgleichungen müssen zusätzlich die Anfangs-/Randbedingungen bzw. der Anfangszustand angegeben werden.
Im 3. Abschnitt sind Pointer auf die Dokumentation der jeweiligen Teilprozesse, die in einer niedrigeren Stufe der Aggregationshierarchie stehen, aufgelistet. Dabei ist zu berücksichtigen, daß die Konsistenz der Umgebungen der zusammengehörenden Realisationen in den verschiedenen Ebenen gewährleistet ist.
Der 4. Abschnitt enthält Pointer auf die zu diesem Prozesstyp gespeicherten Realisationen.

2.2 Dokumentation einer Prozess-Realisation

In der Dokumentation der Prozess-Realisationen sind zunächst die Werte der Parameter zu spezifizieren sowie die Eigenschaften der Prozessumgebung zu charakterisieren. Zum einen ist vorgesehen diese Charakterisierung mit Hilfe von Schlagwörtern aus hierarchischen organisierten Thesauri zu ermöglichen (z.B. Typ des Ökosystems, Klimatyp u.a). Zum anderen ist aber auch die Möglichkeit gegeben spezifische Sachverhalte zu beschreiben, da unter Umständen mit vorgegebenen Thesauri eine ausreichende Charakterisierung der Prozessumgebung nicht möglich ist.
Zur Dokumentation der Prozess-Realisation gehört auch die Angabe der Definitions- und Wertebereiche der verwendeten Variablen.
Ferner sind der Name und Adresse des Modellentwicklers, das ursprüngliche Anwendungsziel sowie Literaturhinweise anzugeben.
Es besteht auch die Möglichkeit, Anwenderkommentare an dieser Stelle zu erfassen und dem Benutzer Hinweise auf andere Modelldokumentationsdatenbanken an die Hand zu geben.

Anhang:

Remark: two different formats are used for mathematical entries:

eq-format: (equation format)
string; (the equation)
one line with max. 80 chars
(brief description of the origin of the equation resp. of the initial/boundary condition)

quant-format: (quantity format)
string (name)
string (unit)
one line with max. 80 chars
(brief description of the quantity)

DESCRIPTION OF A PROCESS

**** #typ.tit **** # title
name of ecological process
Format: one line with max. 80 chars

**** #typ.key **** # list of keywords
list of keywords
Format: list of words; delimiter = TAB

**** #typ.abs **** # abstract
brief description
Format: free text

**** #typ.desc **** # documentation of the process
1. DECLARATION:

1.1 Input:
quant-format

1.2 State:
quant-format

1.3 Parameter:
quant-format

2. EQUATIONS

2.1.1 PDE's
eq-format

2.1.2 initial/boundary conditions for PDE's
eq-format

2.2. Dynamic state equations

2.2.1 type of equation
Format: type =[i,j,k]
i= 0 time continous, i= 1 time discrete
j= 0 state continous, j= 1 state discrete
k= 0 deterministic, k= 1 stochastic

2.2.2 equation
eq-format
2.2.3 initial conditions
eq-format
2.3 algebraic equations
eq-format

3. LIST OF EMBEDDED PROCESSES (SUBSYSTEMS)

f_i: #typ_i.#obj_j => #typ_k.#obj_l, ...
f_i: #typ_i.#obj_k => #typ_o.#obj_p, ...

f_i	= name (symbol) for embedded process
#typ_i	= pointer to the documentation of **this** process-typ
#typ_i.#$obj_{[j,k]}$	= pointer to the documentations of realizations of **this** pocess-typ
#$typ_{[k,o]}$.#$obj_{[l,p]}$	= pointer to the documentations of realisations of **the embedded** process-typ

Format: string: pointer => list of pointers

4. LIST OF REALIZATIONS

#typ.#obj,#typ.#obj ....
Format: list of pointers

DESCRIPTION OF A PROCESS-REALIZATION

**** #typ.#obj.val **** # values
Parameter:
name value [or range]
Format: string number [- number]

**** #typ.#obj.env **** # specific conditions of environment
1. Common classification of environment and system components:
Format:

Type of ecosystem	*: string*
Geographic Location	*: string*
Climatic type	*: string*
Biological classification	*: string*
Soil type	*: string*

2. Specific conditions of environment:

e.g. temperature, pH, ... etc.
Format: free text

**** #typ.#obj.cons **** # range of validty
1. Specification of the domain

1.1 Input:
name lower_limit upper_limit
Format: name number number

2. Specification of the range

2.1 States:
name lower_limit upper_limit
Format: name number number

3. Specification of valid time- resp. time-space-domain

tmax value [unit]
Format: tmax number [string]
xmin xmax xunit ymin ymax yunit
Format: number number string number number string

**** #typ.#obj.aut **** # author(s)
Format:
Name : string
first name : string
Institution/Company : string
Department : string
Postal adress : string
Phone : string
Fax : string
E-mail : string

**** #typ.#obj.sour **** # references
Format:
Author(s) : string
Title : string
Source : string
Publication year : string
Page(s) : string

**** #typ.#obj.appl **** # original application
Format: free text

**** #typ.#obj.comm **** # collection of comments
(mail-)comment
Format: free text

**** #typ.#obj.point **** # pointer to UFIS and Kiel database
Format: list of pointers