

**MANAGEMENTUL ȘI MODELAREA
DATELOR EXPERIMENTALE**

**MASTERAT
BIOLOGIE APLICATĂ I**

**DEFINIȚIA, ALGORITMUL,
CLASIFICAREA,
EVALUAREA ȘI UTILIZAREA
MODELELOR**

Conf. dr. Ioan Sîrbu

TEMATICA:

- Definiții ale modelelor;
- Principiile și algoritmul modelării;
- Categoriile de modele;
- Modele conceptuale;
- Trecerea de la modelul conceptual la cel cantitativ;
- Modele statistice și modele teoretice;
- Evaluarea și utilizarea modelelor.

În conformitate cu dicționarul Oxford **modelul** este o reprezentare a realității, care prezintă în manieră simplificată principalele trăsături ale unor aspecte ale lumii reale, pentru a le face mai ușor de înțeles și, adesea, a facilita avansarea de prognoze (Allaby, 2005).

În conformitate cu DEX al limbii române un model înseamnă o reprezentare simplificată a unui proces sau a unui sistem; este un sistem teoretic sau material, cu ajutorul căruia pot fi studiate indirect proprietățile și transformările altui sistem, mai complex, cu care modelul prezintă analogii.

1. PRINCIPIILE ȘI ALGORITMUL MODELĂRII

Modelul = o reprezentare a unei entități reale, a unui lucru, idei sau condiții, care descrie și sintetizează structura, dinamica sau funcțiile unui sistem sau ale unui proces.

Procesul de modelare = un șir de etape, care trebuie parcurse, pentru a converti o idee, mai întâi într-un model conceptual, și apoi într-o reprezentare calitativă și/sau cantitativă.

Modelarea este o activitate universală

Toți practicienii, uneori și teoreticienii, creează și testează modele de cele mai diferite forme, idei și expresii.

Dacă excludeți modelul dintr-o cercetare, veți fi niște simpli colecționari de fapte sau de date primare.

De ce modelăm?

Modelele:

- orientează, asistă și susțin studiile;
- testează ipotezele de lucru;
- permit înțelegerea sistemelor și proceselor complexe;
- descriu, caracterizează și permit analiza proprietăților emergente ale sistemelor;
- susțin deontologia profesională;
- permit simularea și prognoza;
- permit experimentare în spațiu și timp virtual.

- Pentru că știm și putem!

Definirea modelului:

- grafic (schiță, desen etc.)
- informal (fără formule)
- formal (cu formule matematice)

- Modelele sunt construite pentru a testa, verifica, simula, prognoza, pentru a ne ajuta să înțelegem mai bine cum funcționează procesele și sistemele vii.
- Modelarea nu este un pas final al unui program experimental, ci este o parte integrantă a cercetării.

◆ ALGORITMUL MODELĂRII

- Definirea problemei:

- scop, obiective, delimitare.

Dacă nu știm ce vrem de la modelul pe care dorim să-l realizăm, foarte probabil nu vom obține nici un model.

- Evaluarea complexității :

- adecvată sistemului și problemei;
- alegerea judicioasă a scărilor (spațială temporală, tematică);

Complexitate = număr de variabile măsurate, nr. de ecuații, parametrii etc.

Legea lui D.B. Lee (1973):

“Calculatoare mai mari, suportă greșeli mai mari”
(modele mai complexe produc erori mai grave)

Comentariul lui Silwert (2001):

“S-au cheltuit în decursul timpului milioane de dolari,
pentru a se demonstra valabilitatea legii lui Lee”.

*Un model mai complex nu este în mod necesar mai bun sau mai util.
Modelele trebuie păstrate în variante simple, până când apare
evidența sau necesitatea pentru a le complica!*

Conceptualizarea = etapa în care transpunem problema particulară într-o imagine sau o idee de ansamblu.

Transpunerea grafică a ideilor, a etapelor și vizualizarea relațiilor între diferitele categorii de variabile, formează **modelul conceptual**.

Prin adăugarea de valori, evaluări cantitative ale proceselor și elaborarea de ecuații, obținem modelul cantitativ.

Construirea modelului se referă la alcătuirea acestuia, în termeni de idei, concepte, expresii fizice, ecuații matematice, respectiv programe pe calculator.

Modelele se construiesc în două feluri:

- bază - vârf
- vârf - bază

Modele bază - vârf:

- asamblarea de multe submodele
(multe elemente sau procese, eventual toate)
- tind să fie foarte complicate
- bune pentru înțelegerea funcționării sistemului
- nu dau rezultate prea bune în scopul pentru au fost create
- de obicei sunt modele teoretice (cauzale, explicative)

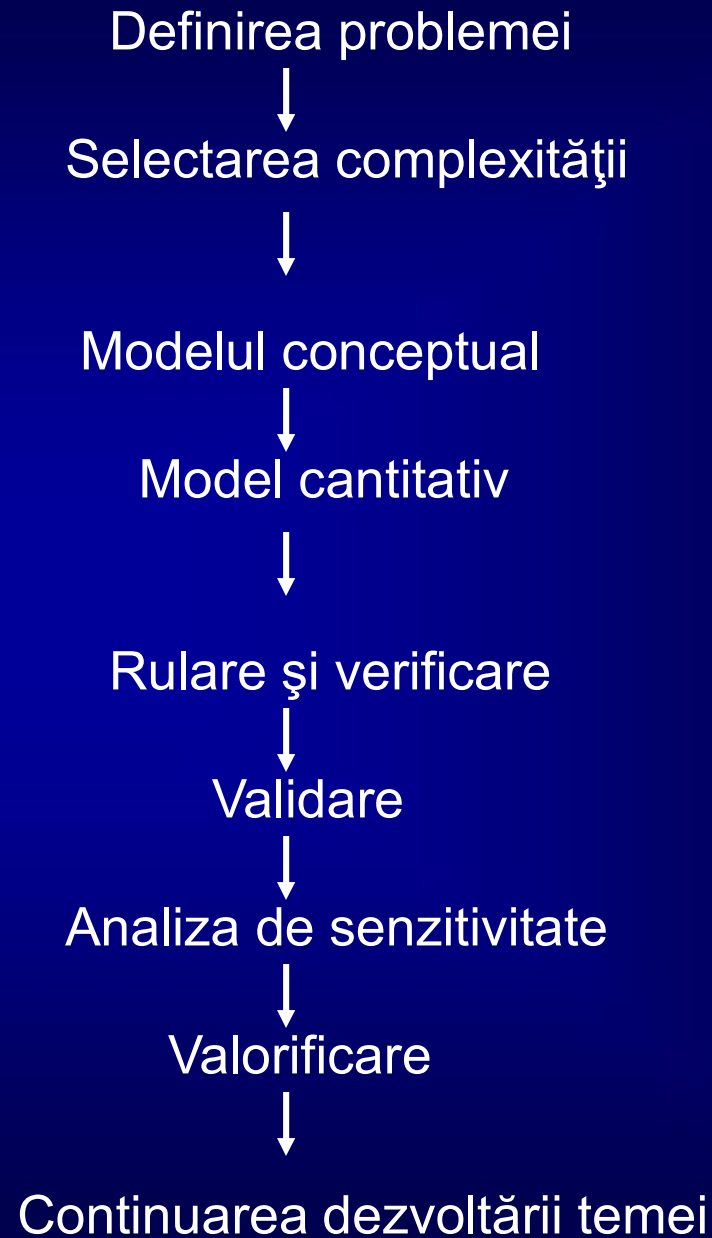
Modele vârf - bază

- pleacă de la funcții generale
(de la comportamentul întregului sistem sau de la proprietățile emergente)
- cele mai multe sunt de tip statistic
- adesea sunt modele empirice
(se bazează pe date experimentale)
- nu presupun considerarea mecanismelor intime
- sunt adesea restricționate la subiectul particular analizat.

Câțiva termeni de specialitate:

- Rularea modelului (funcționare repetată, aplicare de mai multe ori, de exemplu pe calculator)
- Verificarea și depanarea modelului (testarea și îmbunătățirea logicii interne)
- Validarea modelului (evaluarea rezultatelor, corespondența datelor prognozate cu cele reale)
- Analiza sensibilității ("*ce-ar fi dacă ...*", adică schimbarea datelor inițiale, sau utilizarea altor variabile, ecuații etc. și evaluarea rezultatelor în lumina modificărilor)

Schema algoritmului procesului de modelare



Nu există corelații directe între:

- valoarea modelului și complexitatea acestuia,
- scară (spațială, tematică etc.) și complexitate,
- utilitate și adevăr.

"Modelul nu este identic cu adevărul; modelul este minciuna prin care sesizăm și descoperim adevărul"

- Modelarea este un proces iterativ!
- Modelele se dezvoltă și evoluează!
- Un model mai bun =
 - simulare mai realistă
 - prognoze mai bune
 - rezultate mai bine validate
 - unealtă mai performantă
 - realism mai convingător

- Nu există "modele universale"
- Modelele trebuie să fie folositoare
- Modelele trebuie valorificate
- Rezultatele să fie clar prezentate

- Ce software utilizam?

- a. Criterii subiective:

- *cel mai bun software este cel pe care îl cunoști!*
 - *poate este mai bine să operezi decât să programezi!*
 - *dacă cereți ajutor calificat, explicați clar ceea ce vreți să obțineți (traduceți din biologie în..)*

b. Criterii obiective: ce vrem de la model, de la software și unde vrem să ajungem?

Variante:

i. Limbaje generale de programare:

ex. C ++, Visual Basic, Pascal etc.

ii. Programe grafice cu opțiuni simple și rigide:

ex. STELLA, SimuLink, ModelMaker

iii. "Calea de mijloc": programe cu opțiuni largi, includ multe funcții și proceduri, evită detalii tehnice, permit controlul, multe posibilități alternative:

ex. MATHCAD, SYSTAT, STATISTICA!

NOI ALEGEM CALEA DE MIJLOC!

ELEMENTELE CONSTITUENTE ALE MODELELOR:

I. VARIABILE ȘI FUNCȚII EXTERNE

II. VARIABILE DE STARE

III. ECUAȚII MATEMATICE

IV. PARAMETRI

V. CONSTANTE UNIVERSALE

2. CLASIFICAREA

MODELELOR

Modelele pot fi clasificate în funcție de abordare, alcătuire, scop și tipul datelor și rezultatelor.

Categoria de modele	Caracteristici
Științifice Management	- scopul este legat de explicare și cunoaștere - scop de gestiune și control a unui proces sau sistem
Formale Informale	- (de obicei) exprimate prin ecuații matematice - nu sunt exprimate prin ecuații
Deterministe Stocastice	- depind numai de datele de intrare - rezultatele depind de distribuții probabilistice sau de factori aleatori

<p>Reducționiste Holiste</p>	<ul style="list-style-type: none"> - includ cât mai multe detalii și subsisteme - utilizează principii și funcții generale, emergente ale sistemului ca întreg
<p>Statice Dinamice</p>	<ul style="list-style-type: none"> - independente de timp - variabilele sunt funcții de timp, eventual și spațiu sau alți parametri independenți care variază
<p>Dinamice în timp continuu Dinamice în timp discret</p>	<ul style="list-style-type: none"> - urmăresc procesul fără întrerupere în timp, fiind modelate cel mai adesea prin ecuații diferențiale - sistemul se studiază la anumite momente de timp, care alcătuiesc un șir discret; ecuațiile utilizate de obicei fiind de tip recurent.
<p>Dinamice distributive Dinamice integrate</p>	<ul style="list-style-type: none"> - definite prin modificarea graduală a mai multor variabile, adesea independente, exprimate prin ecuații diferențiale parțiale - definite la scară temporală lungă, exprimate prin valori medii ale variabilelor de stare și proces

<p>Liniare Neliniare</p>	<p>- sunt considerate ecuații de gradul I - ecuații de grad superior sau funcții neliniare</p>
<p>Cauzale (cutie albă) Necauzale (cutie neagră)</p>	<p>- datele de intrare și rezultatele sunt conectate prin relații de tip cauză - efect - sunt comparate intrările cu ieșirile, fără analiza mecanismului intim cauzal</p>
<p>Populaționale Suprapopulaționale (Cenotice) Ecosistemice (Holiste) Modele ale bio(eco)sferei</p>	<p>Modele care descriu structura și/sau funcționarea sistemelor supraindividuale la diferite niveluri de organizare (populație, comunitate sau asociație, biocenoză, ecosistem, biom, biosferă respectiv ecosferă etc., supra- și subdiviziuni ale acestora).</p>

Bioenergetice	- descriu și caracterizează fluxul energetic
Biogeochimice	- circuitul materiei sau cicluri ale unor elemente și substanțe în natură
Modele bazate pe indivizi (MBI)	- analizează comportamentele elementelor constituente ca bază explicativă pentru funcționarea întregului;
Agregative	- grupează elementele constituente după anumite criterii.

Alte categorii:

- Modele spațiale
- Modele orientate pe obiecte
- Rețele neuronale
- Modele fractale
- Modele tip Monte Carlo
- Modele de optimizare
- Modele bazate pe teoria haosului
- Modele bazate pe teoria jocurilor
- ...etc.

Elemente ale modelelor dinamice:

- traiectoria sistemului
- starea staționară a sistemului
(*steady state*)
- echilibrul sistemului
- cicluri limită

3. MODELUL CONCEPTUAL

Procesul de conceptualizare =

trecerea de la o idee la expresia imaginativă a acesteia

- parte integrantă a oricărui studiu
- sunt de obicei scheme sau diagrame
- de exemplu conțin blocuri ...
(etape, sarcini, comenzi, întrebări, declarații etc.)
- ... între care amplasăm linii sau săgeți
(relații, legături, interacțiuni, ordinea de parcurgere etc.)

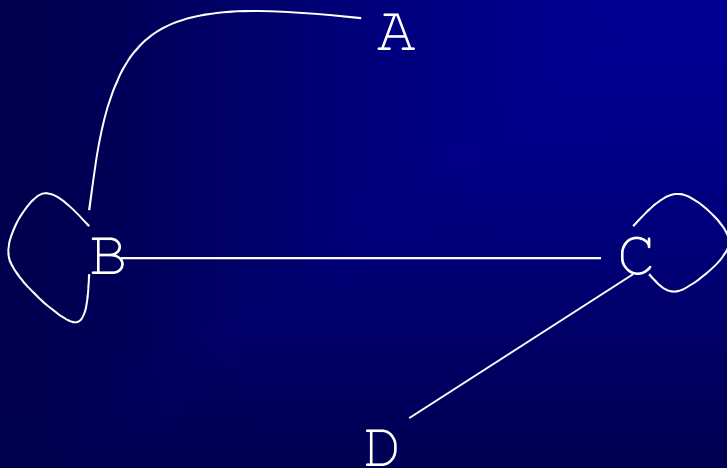
Categorii de modele conceptuale:

1. Modele verbale
2. Scheme și desene
3. Diagrame în blocuri (tip cutie):
 - cutie neagră
 - cutie albă
 - cutie gri
4. Modele I/O (Input/ Output)
5. Modele matriciale

Exemplu: o comunitate formată din 4 specii (A, B, C, D) între care pot exista relații (1) sau nu (0):

Expresie matricială:

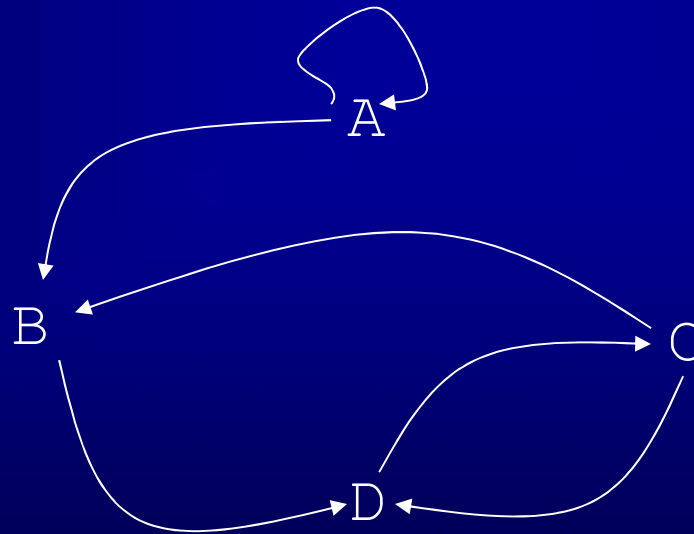
	A	B	C	D
A	0			
B	1	1		
C	0	1	0	
D	0	0	1	0



Expresie grafică
(model conceptual sub
forma unui graf)

Mai complex...

		de la (prădează pe...):			
		A	B	C	D
către: (este prădat de)	A	1	0	0	0
	B	1	0	1	0
	C	0	0	0	1
	D	0	1	1	0

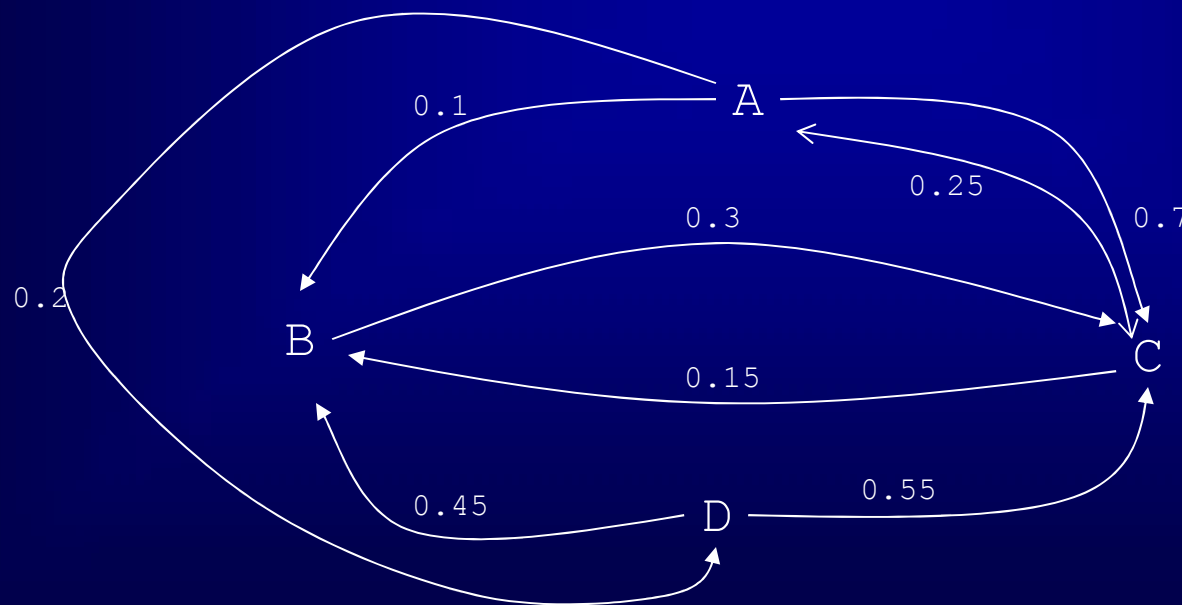


Digraf orientat

Și mai complex...

Coeficienți de competiție:

	A	B	C	D
A	0	0	0.25	0
B	0.1	0	0.15	0.45
C	0.7	0.3	0	0.55
D	0.2	0	0	0



Digraf,
numeric,
orientat

4. TRECEREA DE LA MODELUL CONCEPTUAL LA MODELUL CANTITATIV

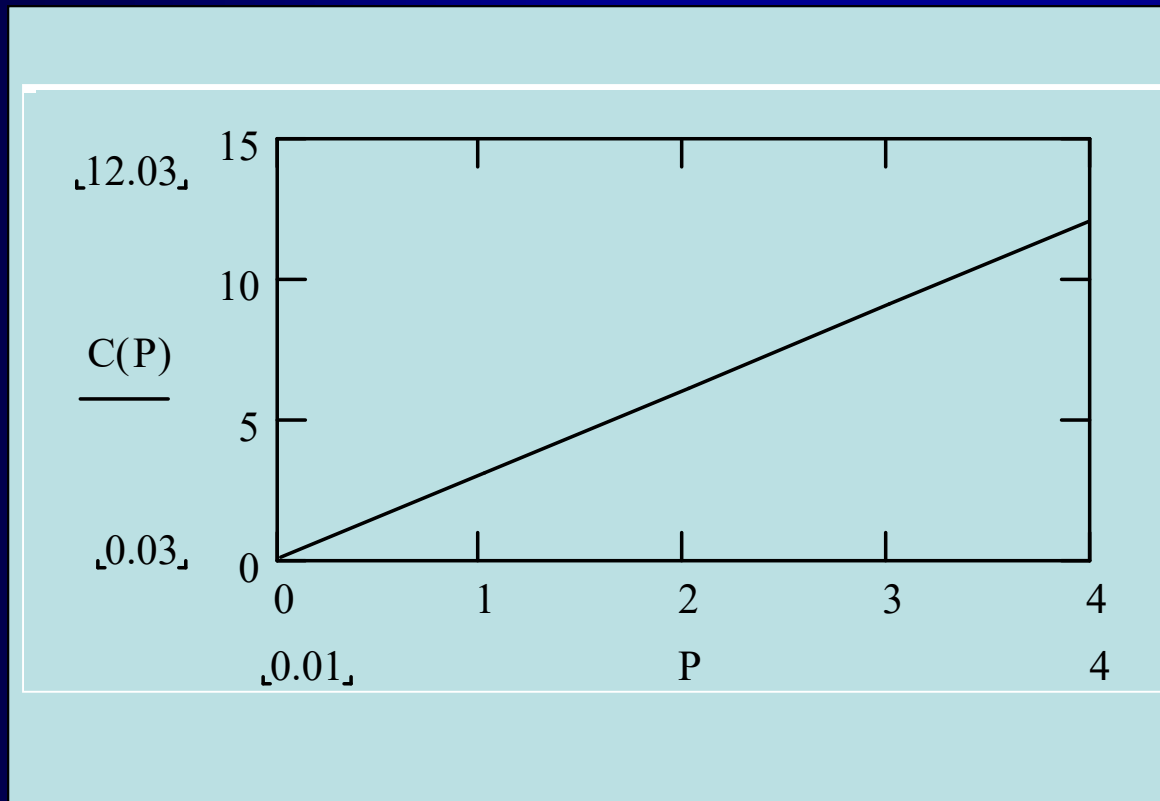
- Un model cantitativ este un set de expresii matematice și valori atașate blocurilor și săgeților modelelor conceptuale.
- Cele mai multe categorii de modele prezentate la clasificare, sunt totodată și cantitative.

Exemplu:

C = rata de prădare (variabila independentă)

P = disponibilitatea prăzii (variabila dependentă)

Trei modele = trei ipoteze

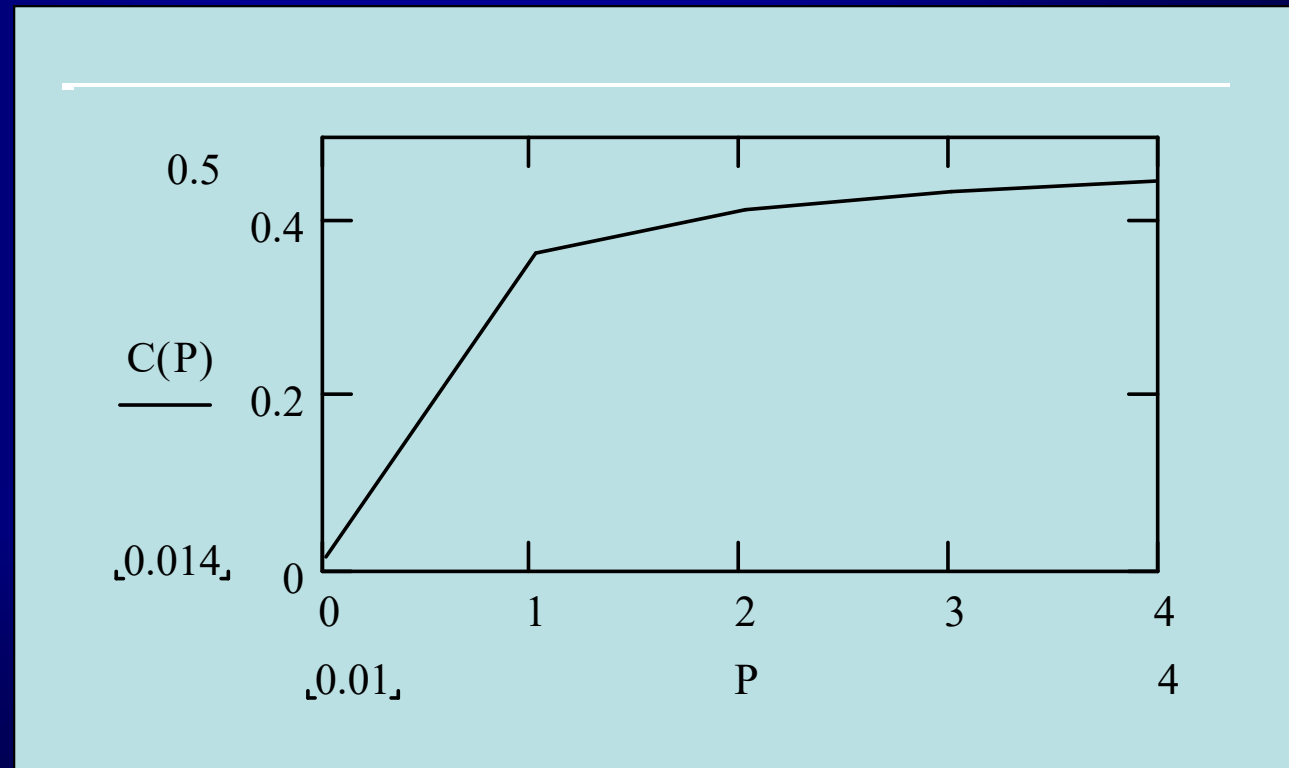


Modelul 1

$$C = a * P$$

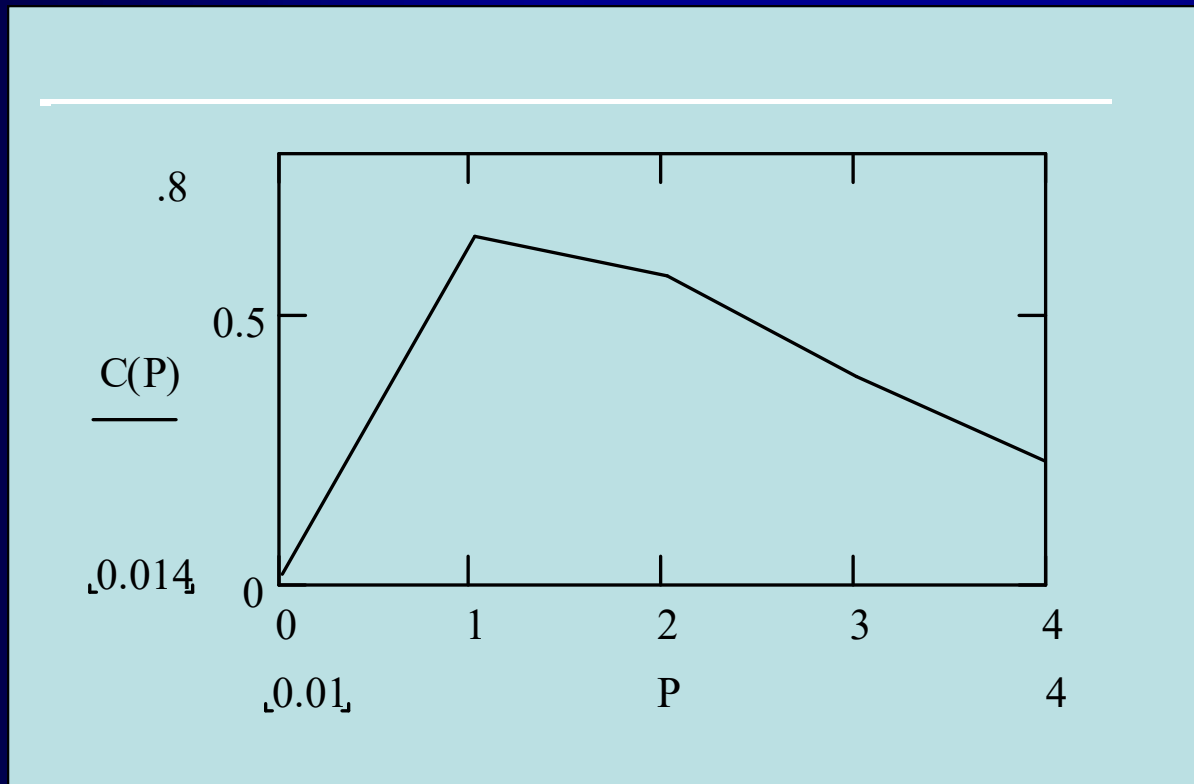
Modelul 2

$$C = (b * P) / (1 + d * P)$$



Modelul 3

$$C = d * P * e^{-j*P}$$



Alți doi termeni (concepte) de specialitate:

- **CALIBRAREA MODELULUI**

(dezvoltarea treptată a modelului și modificarea acestuia, în paralel cu desfășurarea studiului, până la potrivirea datelor prognozate de model cu cele reale, obținute prin experimente sau evaluări)

- **DETECȚIA ȘTIINȚIFICĂ**

(metodă și procedură prin care, pornind de la model, simulând și experimentând pe acesta, deducem legități și explicații ale fenomenelor naturale)

5. EVALUAREA ȘI UTILIZAREA MODELELOR

Evaluarea modelului

- relevarea semnificației rezultatelor;
 - verificarea și testarea logicii interne;
 - testarea comportamentului modelului;
 - potrivirea datelor prognozate cu cele reale;
 - analiza de senzitivitate;
 - relevarea utilității acestuia.
-
- Utilizarea, exploatarea și comunicarea modelului

Evaluarea se referă la verificarea utilității modelului de a deservi un scop pentru care a fost creat.

Nu există modele universale!

În general, un model care este extrem de util pentru un anumit scop, devine repede inutil sau fals pentru alte condiții sau scopuri.

În evaluarea modelelor se consideră o varietate de aspecte și trăsături calitative și cantitative ale acestora, atât în ceea ce privește structura, rezultatele, cât și comportamentul lor.

5.1. Aflarea și testarea soluției (soluțiilor)

- Construirea modelului conceptual;
- Construirea ecuațiilor;
- Soluții și verificări ale rezultatelor;
- Cel mai simplu este să preluăm din literatură rezultatele modelelor cunoscute, consacrate, dacă acestea există.
- Generarea soluției numerice pentru valori inițiale cunoscute, este adesea mult mai facilă decât cunoașterea și evaluarea soluției analitice; eventual o bună variantă este să le aflăm pe ambele și să comparăm rezultatele.

5. 2. Testarea logicii interne a modelului

- Este rezonabil modelul?
- Variabilele de stare rămân pozitive?
- Modelul se comportă așa cum este de așteptat?

5. 3. Validarea modelului

În timpul validării testăm dacă modelul este capabil să reproducă datele reale, atât calitativ cât și cantitativ.

Prin aceasta ne confruntăm cu consecințele modelului.

Dacă produsul modelului este similar cu observațiile, înseamnă că modelul este validat pentru moment.

Aceasta nu înseamnă că modelul este corect, ci doar că nu l-am dovedit a fi fals.

În caz contrar, trei categorii de acțiuni pot fi desfășurate: suspectăm datele, modelul sau parametrii.

- În primul rând se verifică precizia și acuratețea datelor. Acestea pot fi testate prin tehnici de tip AC/CC (asigurarea calității și controlul calității datelor), metode standardizate de colectare și prelucrare statistică a datelor.
- În al doilea rând trebuie revizuit modelul, în ceea ce privește presupunerile, relațiile, complexitatea, ecuațiile. Nepotrivirea dintre date și model poate releva goluri în înțelegerea noastră referitoare la proces sau la sistem.
- În sfârșit poate ar trebui reconsiderate valorile parametrilor (a coeficienților care fac legătura dintre datele reale și ecuațiile modelului). Acest lucru se poate realiza prin calibrare.

- Verificăm dacă nu a fost depășit domeniul de acțiune al modelului, dacă nu au fost violate constrângerile și limitările.
- Este important ca datele reale pe baza cărora se construiesc unele modele (în general cele statistice), trebuie să fie altele decât cele utilizate în evaluarea lor, în caz contrar este imposibil să realizăm o validare obiectivă.
- Aici se pot intercala în mod judicios și teste statistice care au scopul de a evalua semnificația diferențelor dintre cele două seturi de date.

5. 4. Analiza de senzitivitate a modelului

- În timpul analizei de senzitivitate sunt modificate valorile parametrilor, eventual sunt încercate ecuații alternative și diverse combinații de variabile, pentru a estima și evalua comportamentul și eficiența modelului.

- Analiza de senzitivitate se poate realiza în două feluri:

- **Analiza de senzitivitate globală**

Valorile parametrilor sunt modificate în cadrul unei amplitudini mari, dar plauzibile, și se evaluează efectul acestui demers asupra valorii (valorilor) uneia sau a mai multor variabile. Aceasta ajută la sesizarea mai bună a relațiilor de tip cauză-efect.

- **Analiza de senzitivitate locală**

Se estimează efectul modificării într-o amplitudine restrânsă (frecvent foarte mică) a unuia sau mai multor parametri. Acest fapt se poate face prin funcțiile de senzitivitate. Acestea reprezintă senzitivitatea unei variabile y_i a modelului la modificarea unui parametru p_j .

Formula generală a unei funcții este:

$$S_{ij} = \frac{\Delta p_j}{\Delta y_i} \frac{\partial y_i}{\partial p_j}$$

- Media sensibilității tuturor variabilelor modelului se poate calcula în funcție de fiecare parametru și această informație este utilizată pentru a ierarhiza parametrii modelului, în funcție de efectul pe care îl au asupra rezultatului.
- Aceștia sunt cei asupra cărora trebuie acționat și care trebuie utilizați în procedurile de monitoring și management, precum și cei asupra cărora se vor concentra viitoarele priorități de cercetare.
- O altă variantă este de a modifica fiecare element constituent al modelului, în mod sistematic, și de a analiza semnificația rezultatelor după un algoritm.

- Există posibilitatea de a include în analiză o doză de incertitudine, sub forma unor parametri cu variație aleatoare.

Prin analiza de sensibilitate analizăm mai detaliat nivelul de încredere pe care îl putem avea în capacitatea modelului de a răspunde la o întrebare sau de a rezolva o anumită problemă.

5. 5. Utilizarea modelului

- Scopul acestei etape este de a atinge obiectivele pe care le-am fixat la începutul activității de modelare.
- Schema generală a utilizării modelului seamănă foarte mult cu cea a designului experimental în cercetarea empirică.
- La început se dezvoltă și execută un design experimental pentru simulare.
- Apoi se interpretează rezultatele simulării, prin utilizarea aceluiași unelte analitice și scheme interpretative, pe care le folosim și în evaluarea rezultatelor de laborator sau de teren.
- În sfârșit, aceste rezultate trebuie comunicate, adică făcute cunoscute potențialilor interesați, beneficiari, sau comunității specialiștilor.

- **Dezvoltarea și executarea unui design experimental pentru simulări**

Principiile designului experimental pe care le aplicăm în orice cercetare, în teren sau laborator, le implementăm și la nivelul experimentării sau simulării, respectiv a utilizării modelului pe calculator.

Presupunerile realizate a priori în faza de conceptualizare, ar trebui să asigure un ghid pentru designul experimental, deoarece acestea reprezintă ipotezele pe care le testăm.

Dacă modelul conține elemente stocastice, este necesară specificarea numărului de replicări (rulări, colectări de rezultate), care să permită obținerea unor rezultate statistic semnificative.

- Rulăm modelul cu ceea ce știm despre sistem, sau în condiții care corespund la ceea ce am numi "probă martor", și apoi, în mod adecvat, vom simula diferitele tratamente.
- Vom modifica parametrii și valorile variabilelor după principii de tipul blocurilor, a designului factorial, ierarhic etc. adecvat condițiilor de definire a problemei abordate.
- Vom calcula numărul necesar de replicări pentru fiecare analiză ca și cum ar fi probe din mediul real.

- **Analiza și interpretarea rezultatelor simulării**

- Stabilirea ipotezelor, alegerea adecvată a testelor, calcularea parametrilor statistici, decizii și interpretarea rezultatelor.

- În primul rând se compară datele prognozate de model cu cele din mediul real (observații, determinări etc.).

- Comparațiile seturilor de date dintre diferitele condiții de simulare se va realiza prin teste adecvate categoriei de date și a strategiei de design experimental ales.

- Aceleași restricții și reguli care se aplică în testarea statistică a datelor experimentale, sunt valabile și în analiza rezultatelor diverselor simulări.

- **Comunicarea rezultatelor și valorificarea modelului**

Modelele cel mai adesea nu sunt, și nu trebuie să fie, un scop în sine.

Ele trebuie să folosească la ceva, și acel ceva trebuie valorificat.

Valorificare înseamnă explicație științifică, rapoarte de cercetare, lucrări, propuneri de management, decizii politice, soluții economice etc.

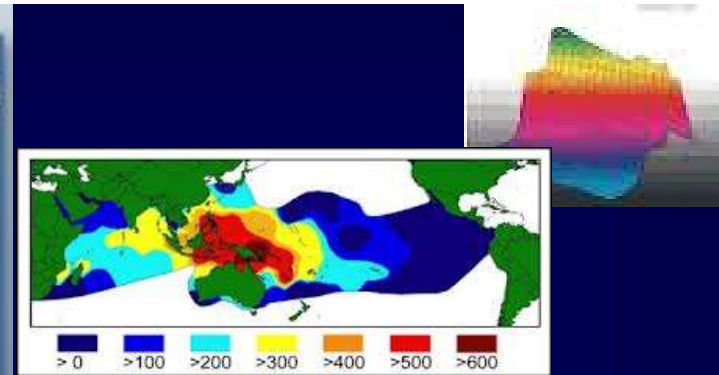
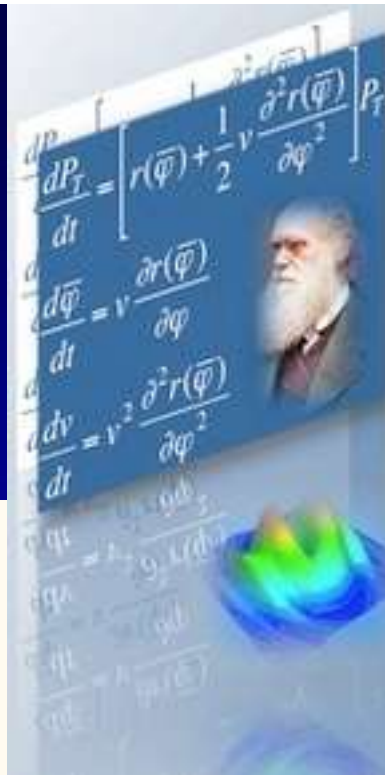
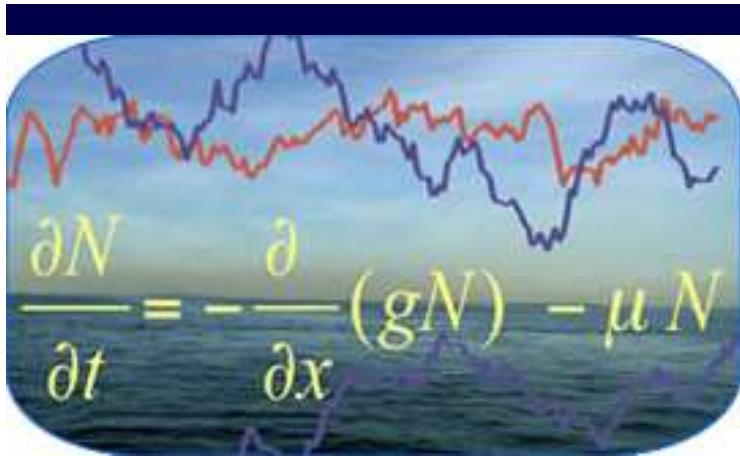
Modelele trebuie făcute cunoscute, ele sunt obiect al comunicării.

La fel cu celelalte etape ale acestei faze, comunicarea nu diferă de principiile și algoritmi care facilitează sau ghidează procesul de elaborare a lucrărilor de specialitate în biologie.

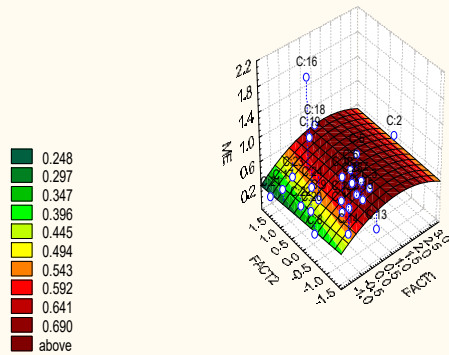
Trebuie să descriem în mod explicit și clar problema, să o punem în context tematic și istoric, să specificăm obiectivele, baza informațională pe care construim studiul, metoda și tehnica folosită (aici se include descrierea formală a modelului, aspecte conceptuale, matematice, informatice, softuri utilizate etc.), rezultatele procedurilor de evaluare și a experimentelor simulate, limitările sau domeniul de valabilitate ale acestora, interpretări și discuții adecvate, precum și concluziile la care am ajuns.

Frecvent este mult mai important să comunicăm și să interpretăm tendințe ale fenomenelor sau proceselor analizate, mai degrabă decât valorile absolute, prognozate, ale acestora.

Prea mulți cercetători devotați analizei și modelării au caracteristica de a intra mult prea mult în detalii, adesea tehnice și relevante doar pentru structura modelului, în loc să se concentreze asupra rezultatelor și semnificației, precum și a prognozelor generale care decurg din utilizarea acestuia.



Model: $me=a+b*fact1+c*fact2+d*fact1*fact1+w*fact1*fact2+v*fac$
 $z=(0.6211085)+(0.1412512)*x+(-0.0318823)*y+(-0.0524606)*x*x+(0.0127711)*x*y+(-0.000781)*y*y$



3D Surface
 $z=2.098+0.981*x-0.132*y-0.218*x*x-0.396*y*y+0.062*y*y$

