

## INTRODUCERE ÎN MODELARE: ALGORITMUL ȘI PRINCIPIILE MODELĂRII

**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 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.

Acestea pot fi foarte de simple, de exemplu o expresie verbală cu privire la însușirea unui subiect, sau o diagramă, cum ar fi două dreptunghiuri înzestrate cu o anumită semnificație, legate între ele printr-o săgeată, care indică o relație. La cealaltă extremă, modelele pot fi oricât de complexe, de exemplu programe imense care rulează pe rețele de calculatoare interconectate, la care lucrează instituții cu sute de oameni, din cele mai diverse specializări. **Procesul de modelare** este un șir de etape, care trebuie parcurse, pentru a converti o idee, mai întâi într-un model conceptual și apoi într-o reprezentare cantitativă.

**Modelarea este o procedură care însoțește inevitabil orice studiu.** Aceasta nu înseamnă neapărat calcul, dar cel mai frecvent implică sau se exprimă (și) prin limbaj matematic, respectiv reclamă utilizarea adecvată a tehnologiei informației. 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*** (W. Silvert, 2001).

**Modelarea** este încadrată variabil, undeva între o simplă și frecvent ignorată metodă a științei, și (polul opus) un domeniu de vârf care necesită specialiști “modelatori”, definindu-se astfel o disciplină distinctă. Adevărul este undeva pe la mijloc.

Modelele sunt pretutindeni în viața noastră. De multe ori când mergem la cumpărături, avem o listă de produse și cantitățile acestora scrise, eventual și prețurile, pentru a nu depăși bugetul alocat sau disponibil. Lista este un model (dacă avem și prețurile este chiar un model cantitativ). Acest model ne ajută să organizăm și să optimizăm ceea ce facem, să ne încadrăm în buget și să alegem judicios produsele necesare. De ce este un model? Pentru că, de exemplu, cuvântul ”unt” este o reprezentare, un cod, și nu poate fi întins pe pâine. La fel și prețul acestuia: nu este totuna să-l aveți scris sau să aveți banii corespunzători în buzunar.

Când planificăm și desfășurăm un studiu, o listă de necesar și etape care trebuie parcurse, respectiv un model, ne ajută într-o manieră similară. Laboratorul și experimentele pe care le desfășurăm acolo, sunt modele ale naturii. Datele, managementul acestora, analiza, reprezentarea rezultatelor, testarea ipotezelor și interpretarea, toate acestea sunt, sau se bazează pe, modele. Testarea ideilor, testarea ipotezelor, înseamnă construirea și testarea modelelor. Un grafic este un model. Grafice de tip PCA sau RDA, sunt chiar modele complexe, multivariate, cantitative, explicative și testabile. Deja ați lucrat cu asemenea unelte. La fel și cuvintele: sunt modele ale obiectelor sau ideilor.

Modelarea servește la descrierea, sinteza și explicarea fenomenelor, este un aparat științific prin care distilăm principalele trăsături emergente ale proceselor și sistemelor, pe care le analizăm, le explicăm și pe care experimentăm în spațiul virtual. Aceasta este indispensabilă pentru orientarea și planificarea studiilor, testarea ipotezelor multiple de lucru, simulare și prognoză, dar - mai ales - ca unealtă eficientă în management și luarea deciziilor.

Cu toate aceste afirmații, trebuie înțeles că modelele nu sunt, și nu trebuie să fie, un rezultat final al cercetării sau un scop în sine. Modelele sunt și trebuie să fie folositoare, printre altele trebuie să contribuie la cunoaștere și înțelegere, să susțină cercetarea și managementul, să prefigureze noile sarcini care le stau în față. Nu în ultimă instanță ele ne ajută să construim un mediu abstract, virtual, în care să experimentăm și să găsim căi sau soluții rezonabile, cu multă vreme înainte de a interveni asupra sistemului sau a procesului real, prin implementarea unor măsuri mai mult sau mai puțin inspirate. Modelarea ne ajută să ne comportăm în spiritul deontologiei profesionale.

**Modelele pot fi definite în mod grafic, informal** (fără ecuații: de exemplu un mod alternativ, preliminar, este definirea verbală a modelului) **sau formal**, prin expresiile matematice corespunzătoare și programe asociate pe calculator (Silvert, 2001). Ilustrarea grafică a unui model se poate realiza printr-o diagramă conceptuală (de exemplu circulația energiei printr-un nivel trofic, circuitul apei sau al unui element oarecare în natură etc.).

Abordarea clasică, obișnuită în procesul de modelare, începe cu caracterizarea și definirea procesului sau a sistemului, avansarea unor ipoteze despre structură și funcții, traducerea acestora într-un sistem formal (de exemplu prin ecuații) și transpunerea pe calculator, într-un program accesibil, rularea și colectarea rezultatelor. Urmează verificarea potrivirii dintre datele reale și cele avansate de model, și recalibrarea modelului pentru a reduce posibilele nepotriviri dintre valorile experimentale și cele așteptate.

Câteodată acest algoritm este posibil și eficient, însă nu totdeauna, atât din motive obiective (legate de specificul sistemului analizat), cât și subiective (pregătirea cercetătorului, disponibilul tehnic și informațional etc.). De multe ori

pornim de la datele personale sau cele obținute de un alt cercetător și construim singuri ecuațiile modelului. Alteori, este recomandat să pornim de la anumite ecuații prestabilite de alții și să utilizăm datele experimentale pentru calcularea de parametri sintetici sau derivați. Dar, din ce în ce mai des, modelarea nu mai implică lucrul cu ecuații, sau cel puțin nu presupune că cercetătorul este obligat să le definească singur, în termeni matematici. O variantă este să lase calculatorul să se descurce cu ceea ce modelatorul cunoaște numai conceptual sau informal, utilizând programe care permit un oarecare grad de independență, care operează cu proceduri și obiecte relativ independente. Procesul de modelare nu are o singură rețetă și nu se poate reduce la un singur algoritm; totul se adaptează la scopul și obiectivele cercetării, la categoria de studiu în derulare, tipul, calitatea și cantitatea datelor. De aceea mai jos redau mai degrabă un ghid, în sensul unei serii de indicații, și nu o colecție de reguli cu rol de rigidizare sau standardizare a demersului.

**Algoritmul modelării** începe ca oricare alt subiect de cercetare cu **definirea problemei**, etapă crucială a oricărei teme. Definirea scopului, a obiectivelor, respectiv a ceea ce vrem să obținem de la modelul pe care îl construim, trebuie să fie delimitat în termeni de subsisteme, timp și spațiu, precum și adaptat la acele informații pe care suntem în stare să le obținem.

În acest proces primul răspuns pe care îl căutăm este cel la întrebarea: “în care scop construim acest model”? Acesta ne ajută să înțelegem motivul demersului nostru, putem stabili liniile de elaborare a etapelor și ne imaginăm la ce va ajuta în final una alta construită (idem). ***Dacă nu știm ce vrem de la modelul pe care dorim să-l realizăm, foarte probabil nu vom obține nici un model.*** Să nu uităm principiile ale cercetării care au implicație directă și în acest domeniu, dintre care amintesc (parafrază după Krebs, 1989):

- *Nu tot ce poate fi modelat, merită să fie modelat!*

- *Consideră și evaluează numai acele variabile și date care răspund la întrebările puse!*

**Evaluarea complexității** pe care o adoptăm este o altă problemă. Adesea se consideră că un model este cu atât mai bun, cu cât este mai complex, respectiv include mai multe variabile, ecuații și subsisteme (dacă se poate toate cu puțință de decelat sau imaginat). Argumentul practic împotriva unei complexități prea mari este centrat în creșterea potențialului de eroare. Fiecare model include o eroare posibilă (adesea cunoscută, sau cel puțin estimată) pe care încercăm să o menținem la un nivel tolerat. Cu cât sunt mai multe ecuații care aproximează procesele evaluate, cu atât eroarea crește. Complexitatea este direct corelată cu gradul de nedeterminare, dificultatea simulării, și adesea este invers corelată cu relevarea semnificației și a potențialului de verificare a rezultatelor. Jørgensen (1988), Jørgensen și Bendoricchio (2001) sugerează includerea în model exclusiv a factorilor, variabilelor și parametrilor care au o influență majoră sau chiar determinantă asupra subiectului

de interes. Complexitatea depinde și de cantitatea și calitatea informației care este disponibilă. Niciodată creșterea complexității modelului, a numărului de ecuații sau a programului pe calculator, nu trebuie să fie un scop în sine. Celebre sunt, în acest sens, legea lui D.B. Lee din 1973: “calculatoare mai mari, fac greșeli mai mari” (modele mai complexe produc erori mai elevate) precum și comentariul aferent al lui Silvert (2001): “s-au cheltuit în decursul timpului milioane de dolari, pentru a se demonstra valabilitatea legii lui Lee”.

Prin urmare selectarea complexității depinde de tipul problemei, a sistemului sau procesului supus analizei, și a datelor disponibile.

Scara la care definim modelul poate fi oricât de mică sau de mare; nu există scări “bune”, “adevărate”, respectiv nici “false” sau “contraindicate” la modul absolut. Dar este sigur că, un subiect particular se pretează mai bine la a fi analizat la o anumită scară și prea puțin la o alta. De asemenea, nu există nici o relație între scara spațială și complexitatea modelului (Silvert, 2001).

**Conceptualizarea** este etapa în care transpunem problema particulară într-o imagine sau idee de ansamblu, cum ar fi de exemplu sub forma unei diagrame, care conține informații despre parametri și despre relațiile dintre aceștia (matematice sau de altă natură). 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 pot construi în două feluri: **de la bază la vârf sau invers**. Modelul **bază - vârf** se realizează prin asamblarea de submodele, respectiv de cât mai multe elemente constituente sau procese care se desfășoară într-un sistem. Acestea tind să fie foarte complicate, deoarece teoretic totul poate fi inclus; se consideră a fi bune pentru înțelegerea funcționării sistemului, dar nu dau rezultate prea bune în scopul pentru au fost create (W. Silvert, 2001). Modelele **vârf - bază** pleacă de la funcții generale, de la comportamentul întregului sistem sau de la proprietățile emergente. Cele mai multe sunt de tip statistic. Acestea sunt modele empirice, se bazează pe date experimentale, nu presupun obligatoriu considerarea mecanismelor subiacente, iar ca domeniu de valabilitate sunt adesea restricționate la subiectul particular analizat (idem).

Atunci când rezultatele obținute prin model nu verifică datele experimentale, sau - și mai rău - când prognozele nu se adevăresc nici pe departe, cel mai adesea este blamat modelul, dar aceasta nu este singura cauză posibilă. Sunt trei posibilități de fapt: model eronat, date experimental- observaționale greșite sau ambele. Trebuie înțeles faptul că datele obținute prin experimente sau observații sunt supuse unui șir de erori potențiale, ele fiind la rândul lor numai o reprezentare a realității și nici pe

departe realitatea însăși. Într-o altă ordine de idei, matematica poate fi perfectă, dar aparatul de măsură nu a fost calibrat, sau eroarea sistematică diferă la cei care execută măsurătorile.

Dacă am ajuns la etapa implementării acestuia pe calculator, pasul care urmează în mod logic este **rularea modelului** de cât mai multe ori, pentru a se evalua construcția obținută. El se rulează cu o varietate de condiții modificate ale parametrilor de intrare, în scopul urmăririi modului în care se comportă în situații cât mai diverse.

**Validarea și verificarea modelului** se execută în secvența următoare a algoritmului. **Evaluarea rezultatelor** înseamnă **validarea** lor, adică corespondența cu valorile reale (observate sau experimentale, în cazul în care acestea există), dacă sunt logice, posibile, îndeplinesc sau nu ceea ce ne-am putea aștepta de la modelul realizat, precum și dacă le-am interpretat corect. Dacă modelul nu trece de acest examen, este necesară întoarcerea la o etapă anterioară a algoritmului, și reluarea procesului ori de câte ori este necesar. Există o deosebire certă între verificare și validare: primul termen semnifică testarea logicii interne a modelului, constând mai ales dintr-o evaluare subiectivă a modului în care s-a alcătuit și cum se comportă acesta (uneori se vorbește în jargonul de specialitate și de **depanarea modelului**). Validarea se realizează prin compararea rezultatelor modelului cu datele reale, care rezultă din analiza directă a sistemului analizat.

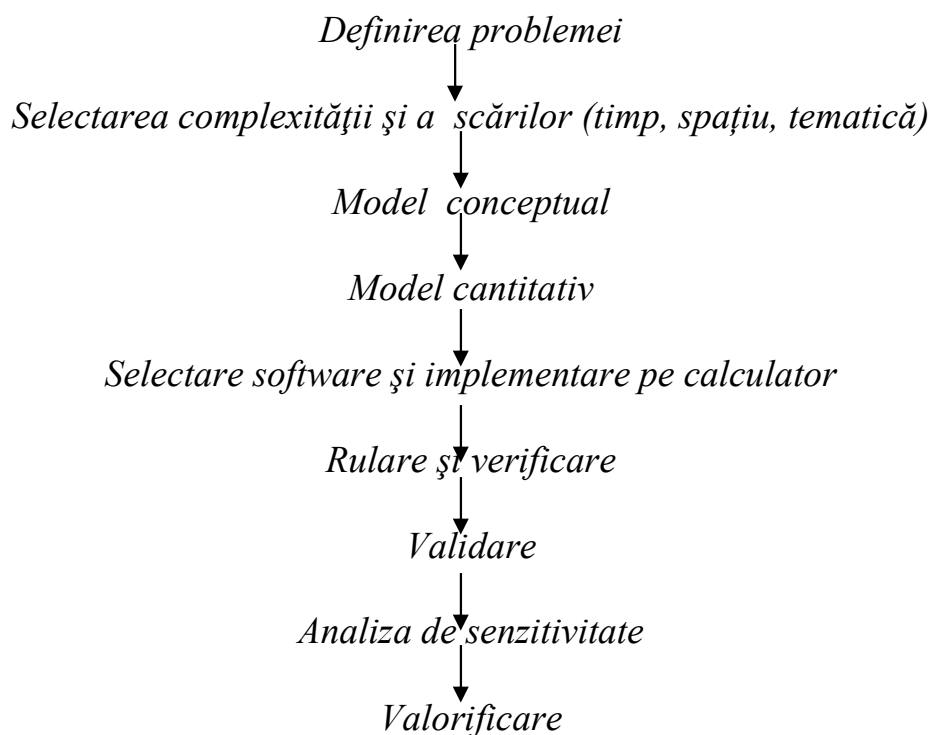


Fig. 1. Algoritmul simplificat al procedurii de modelare

**Un model bun (validat) va fi cel care oferă rezultate, ce se potrivesc cu fenomenele și valorile proceselor reale, și care au o logică științifică.**

Adesea se realizează și **analiza senzitivității**, care investighează cum s-ar modifica concluziile dacă parametrii, datele inițiale și/sau ecuațiile ar fi diferite. Această analiză este utilizată pentru investigarea robusteții modelului, dar și pentru a ghida cercetarea viitoare, respectiv pentru a sesiza importanța diferitelor elemente constituente ale modelului în ceea ce privește contribuția la rezultate.

Am amintit anterior două mituri false ale modelării: relația dintre valoarea modelului și complexitatea acestuia, precum și absența legăturii dintre scara de raportare a sistemului sau procesului și complexitate.

Un alt mit este cel al relației absolute între utilitate și adevăr. Este bine de știut că un model poate să nu fie absolut adevărat (deci valabil în toate condițiile) și totuși să fie extrem de util. De exemplu, legile lui Newton au fost multă vreme considerate infailibile, până în 1905, odată cu apariția teoriei lui Einstein, când au fost contestate. Teoriile newtoniene sunt astăzi recunoscute ca fiind valabile numai pentru obiecte care se mișcă încet și cu energie redusă, dar sunt în continuare mai mult utilizate în viața practică decât teoriile lui Einstein.

În schema algoritmului de modelare (fig. 1) am ales numai etapele principale care se referă la categorii de modele ce implică și formule, respectiv care pot fi transpuse pe calculator. Este evident că alte categorii (modelele grafice, multe modele statistice care presupun operarea cu ajutorul unui program prestabilit etc.) nu includ în mod obligatoriu unele etape ale acestui algoritm, iar altele necesită secvențe suplimentare, inclusiv ramificații condiționate de rezultatele intermediare (de exemplu la modelele stocastice). Dar unele etape, cum ar fi: definirea problemei, crearea imaginii acesteia, evaluarea calității datelor existente, verificarea modelului și validarea rezultatelor fac parte din aproape toate procesele de modelare.

**Modelarea este un proces iterativ.** Dacă un model a fost validat, nu înseamnă că va fi abandonat. Date noi, idei, includerea altor variabile, redeschid de fiecare dată problema dezvoltării și conduc la modificarea perpetuă și evoluția oricărui model. Un model mai bun înseamnă o simulare mai realistă, prognoze mai bune, acuratețe sporită a rezultatelor, uneori includerea mai multor detalii și tendința spre o cât mai bună apropiere de fenomenele naturale. Jørgensen (1988) susține însă că pentru un interval dat de timp, resursele limitate vor opri mai devreme sau mai târziu evoluția oricărui model. O caracteristică a modelării este **absența a ceea ce se numește un model universal** (deci cu valoare generală pentru o categorie de sisteme sau procese). Chiar dacă apar în literatura de specialitate expresii ca: “modelul populației” sau “al lacului” etc., aceasta nu înseamnă că există un singur model, reprezentativ sau ideal, la care se raportează sistemele particulare studiate sau o categorie definitorie. Printr-o categorie de model vom înțelege aspectul central al abordării metodologice (de exemplu modelul matricial se referă la faptul că

abordarea se face prin algebră matricială; modelul diferențial nu implică o anumită ecuație ci o abordare cu un anumit specific matematic a fenomenelor care sunt monitorizate în timp continuu etc.).

Modelele trebuie să fie **folositoare**, adică să fie prea rar un scop în sine. Ele sunt construite pentru a testa, verifica, simula sau prognoza, respectiv pentru a ne ajuta să înțelegem mai bine modul în care funcționează sistemele, și se desfășoară procesele. Modelarea nu este pasul final al unui program experimental, ci este o parte integrată în munca de cercetare.

**Valorificarea modelului** îmbracă numeroase aspecte, similare cu cele care decurg din utilizarea rezultatelor cercetării. Unele modele sunt construite ca aparat științific, ne ajută să înțelegem și să explicăm, iar valorificarea poate însemna publicarea și popularizarea rezultatelor. Altele sunt dezvoltate ca unelte de management, iar valorificarea va consta în implementarea rezultatului în lumea reală, în scop protectiv, ameliorativ sau economic.

**Prezentarea rezultatelor** nu este deloc o etapă de neglijat în complexul finalizării unui studiu. Un model nu este folositor dacă rezultatele sunt ascunse, ignorate sau sunt de neînțeles. Ceea ce am obținut trebuie făcut cunoscut și altora.

#### ◆ Ce software utilizăm?

Este întrebarea care apare mai devreme sau mai târziu în orice studiu cu pretenții, precum și în evoluția capacităților cercetătorului. Este bine să subliniem, că nu există și nu poate exista un singur răspuns la aceasta, deoarece sunt cel puțin două categorii de motivații: de ordin subiectiv și obiectiv. Cele subiective sunt legate de reguli ale cercetării, care afirmă că: *“cel mai bun software este cel pe care îl cunoști”*, apoi *“cea mai bună tehnică de calcul este cea la care ai acces nelimitat”*, precum și *“este preferabil ca un cercetător să învețe să opereze decât să programeze”* (Sîrbu și Benedek, 2012). Este evident că un limbaj sau program pe care nimeni din echipă nu-l știe, nu are nici o șansă să rezolve problema, indiferent cât este de nou sau sofisticat. Snobismul modei dezvoltării neîncetate a calculatorului personal, veșnica vânare a unui echipament mai performant, programe mai diverse și mai noi, a avea permanent ultima placă video etc., sunt fără îndoială utile unui informatician sau unui pasionat al jocurilor, dar sunt preocupări prea puțin potrivite unui biolog practicant. Pe lângă confuzia între scop și mijloace, implică și pierdere de timp și alocarea nejustificată a resurselor. Mult mai folositoare este alocarea de timp pentru cunoașterea aprofundată a modului de utilizare corectă a acelor programe, care într-adevăr ne ajută să progresăm în cercetare și în folosirea rezultatelor, precum și actualizarea lor și a echipamentului hard atunci când este necesar și oportun. Apelarea la un informatician sau (în funcție de natura problemei) un matematician, poate da uneori rezultate, dar atenție la bariera de comunicare și

de reprezentare a intențiilor. Includerea unui specialist din alt domeniu este benefică pentru echipă, numai dacă există dorința de comunicare bidirecțională și traducerea tuturor etapelor în entități cu semnificație similară pentru ambele părți. Orice biolog sau ecolog ar trebui să știe ce înseamnă specializare în cadrul domeniului (deși ecologii, mai ales teoreticienii holiști, tind uneori să uite spre paguba lor aceasta), dar consideră alte discipline ca fiind omogene. Și matematicienii, respectiv informaticienii, sunt specializați, ceea ce nu înseamnă că știu neapărat multe despre puțin, ci faptul că știu mult mai mult despre unele lucruri și semnificativ mai puțin despre oricare altele. Este o mare deosebire între calitatea/cantitatea informațiilor pe care o persoană le deține și potențialul acesteia de a selecta informația relevantă într-un anumit context (marea problemă a celor care predau la alte specializări).

Există și argumente obiective pe care le avem în vedere atunci când selectăm software-ul. Când suntem puși în situația de a alege, ne confruntăm cu o plajă mărginită de două extreme. Astfel, pe de o parte există limbaje de programare generală (de tipul C ++, Visual Basic, Pascal etc.) care permit individului un control complet asupra procesului de modelare, dar care reclamă cunoștințe apreciabile de informatică și matematică. La cealaltă extremă sunt programele grafice de tipul STELLA, Simulink sau ModelMaker, care permit utilizatorului să selecteze din meniuri și galerii de icoane elementele necesare construirii modelului, lăsând calculatorul să se ocupe de toate detaliile. La acestea dezavantajele sunt legate de limitarea potențialului de modelare, a opțiunilor utilizatorului, lipsa de elasticitate a procesului și constrângerea la ceea ce este deja preprogramat. Calea de mijloc (pe care de altfel o adoptăm în studiul modelării la specializarea noastră) o constituie sistemele de programe care includ funcții și proceduri (inclusiv grafice) care permit evitarea multor detalii de ordin informatic și matematic, dar care lasă o gamă largă de posibilități alternative cercetătorului. Printre variantele de acest gen, flexibile și facil de studiat, amintesc softurile de tipul R și Canoco (mai ales pentru analiză și modelare statistică), iar pentru începători uneori folosim și Mystat sau altele, respectiv, mai utilizăm și Mathcad, pentru o serie mai largă și diversă de categorii de modele matematice (de exemplu modelare prin calcul diferențial, modele dinamice în timp discret sau continuu, calcul matricial etc.)



## MODELE STATISTICE ȘI MODELE TEORETICE; EVITAREA ERORILOR DE MODELARE

Statistica este un limbaj al științei, care ne permite să înțelegem natura (Zimmerman, 1995 ap. Franklin et al., 2001).

Lumea este caracterizată prin incertitudine: toate sistemele și procesele prezintă și fluctuații cu caracter imprevizibil, de "zgomot". Ca urmare sistemele observaționale, dintre noi și natură (lumea reală), conțin atât de multe surse de incertitudine, încât chiar dacă procesele sunt deterministe, previzibile, observațiile sunt de obicei nesigure. Dacă înțelegem teoria ca o forță ce trece dincolo de date, care înglobează cazurile particulare în ansambluri integratoare, atunci putem afirma că orice model, indiferent de forma și construcția lui, este o formă a teoriei științifice (Mangel și col., 2001). În această abordare putem identifica o **nouă clasificare a modelelor: teoretice și statistice.**

**Modelele statistice** sunt utilizate pentru obținerea, analiza și interpretarea datelor. Acestea utilizează în special metode statistice inferențiale, cum ar fi de exemplu definirea și estimarea parametrilor pe baza datelor reale. Modelul statistic cel mai comun are un singur parametru populațional - media ( $\mu$ ), pe care o estimăm pe baza parametrului statistic corespunzător (media aritmetică a șirului de date) și a limitelor de confidență (a se revedea Sîrbu și Benedek, 2012). Pe măsură ce înaintăm în studiu, am învățat (sau vom învăța în continuare) să realizăm analize de regresie (univariată și liniară), de corelație, primii pași în testarea ipotezelor, ANOVA, analize statistice multivariate de ordonare (PCA, CA, RDA, CCA etc.), de agregare/clasificare și altele. Toate acestea sunt totodată și modele statistice.

Modelele statistice **nu iau în considerare mecanismele intime ale proceselor; ele stabilesc relații între variabile sau categorii de fenomene, fără analiza cauzalității lor.** După G.C. White (2001) modelele statistice au trei funcții importante. În primul rând ele pot fi testate și verificate prin datele reale. În al doilea rând, dacă le acceptăm ca unelte abstracte utile pentru descrierea fenomenelor naturale, le putem folosi pentru analiza calității datelor reale de care dispunem. În sfârșit utilizarea modelelor și datelor pentru estimarea parametrilor, poate releva funcții și însușiri emergente ale sistemelor, stabili prognoze ale dinamicii proceselor, sau poate orienta managementul (de exemplu al proiectelor dau al resurselor naturale etc.).

Prin urmare modelele statistice sunt de tip "*post-hoc*", adică sunt realizate după ce datele au fost colectate.

Prin contrast, **modelele teoretice (uneori sunt denumite și modele deterministe în sens larg)** includ descrierea mecanismelor, a relațiilor cauzale

dintre fenomene, conducând astfel spre prognoze, înainte ca datele să fie colectate (Mangel și col., 2001). Altfel spus, ele sunt de tip "*ante-hoc*". Când utilizarea modelelor de acest gen nu produce convergența prognozelor cu rezultatele, înseamnă că trebuie să reconsiderăm modelul teoretic, logica acestuia, precum și calitatea datelor disponibile.

Modelele statistice sunt frecvent agregative, adică grupează categoriile de elemente după expresii medii ale însușirilor sau performanțelor lor. Acest fapt îndeamnă la prudență, mai ales atunci când traiectoria sistemelor este influențată de variabilitatea sau comportamentul individual al elementelor constituente. De exemplu, multe modele ale unor însușiri emergente, cum ar fi cele de creștere populațională, energetică comunităților etc., depind explicit de ratele de creștere, mortalitate, natalitate, reproducere, strategii comportamentale active de selectare și exploatare a teritoriului, și multe altele. Aceste însușiri individuale prezintă o relativă autonomie, producând potențial modificări însemnate în rezultatele modelului sistemului care le integrează (idem).

Este benefic ca în procesul de modelare să tratăm separat întrebări de tipul "de ce" (distale/ funcționale), de "cum" (proximale/ mecaniciste). Primele se adresează cauzelor fenomenelor, fiind fundamentate, în cazul sistemelor vii, prin teoria evoluției prin selecție naturală sau artificială. Modelele care se ocupă de întrebări proximale, se concentrează asupra modului în care operează un mecanism și cum se realizează un proces (idem).

Acele modele care sunt, simultan, cele mai explicative (adică ne ajută să înțelegem datele) și totodată cele mai bune predictive (prognozează rezultate înainte de a avea dovezile din lumea reală, ne ajută să identificăm noi date prin extrapolarea rezultatelor viitoarelor experiențe) sunt totodată și cele mai valabile, respectiv utile, pentru scopul în care au fost alcătuite.

După cum am mai afirmat, modelele statistice sunt cel mai frecvent limitate la procesul sau sistemul particular studiat, respectiv sunt valabile pentru seturile disponibile de date experimentale. Ele nu sunt **generalizabile** și nici - un alt termen de specialitate - **transpozabile** (adică nu pot fi aplicate la un alt set de date, decât în eventualitatea unor modificări adecvate a parametrilor, respectiv a recalibrării eficiente). Prin contrast, modelele teoretice, care includ descrierea mecanismelor intime, sunt transpozabile și - în limitele definirii unor anumite seturi de variabile independente, cu domeniile aferente de variație - sunt și generalizabile. Oricum, aceste considerente pot fi supuse și diverselor excepții sau critici; depinde de modul în care definim problema, de datele disponibile, și de cum știm să extrapolăm rezultatele.

- **Evitarea erorilor de modelare**

Criteriile stabilite mai sus ne permit să distingem baze obiective pentru selectarea unor modele din ce în ce mai valabile, dar nu spun nimic despre tehnicile de evitare a erorilor, respectiv despre capcanele care pot apărea în modelarea statistică.

O primă regulă - generală - este aceea că mai multe variabile și - implicit - parametri, acumulează o cantitate mai mare de nedeterminare și de variații accidentale ("zgomote"), prin urmare scad precizia modelului. Altfel spus, acuratețea și precizia modelului va crește până la un anumit număr de variabile și parametri independenți, după care va tinde să scadă accelerat cu numărul acestora. Concluzia este aceea de selectare a unui număr optim, și (preferabil) redus, de parametri, precum și evitarea celor care implică prea multă incertitudine (Mangel și col., 2001). Astfel se aplică cel de-al doilea paradox al lui Ludwig, care mai poate fi enunțat (și) prin principiul imperativ: **"modelele trebuie păstrate în variante simple, până când apar motive serioase să le complicăm"**.

O a doua idee este cea de a dispune totdeauna de **mai multe ipoteze de lucru** pentru analiza și modelarea acelui set de date. Altfel spus, să nu ne limităm niciodată la o singură abordare, respectiv un singur tip de model, ci să încercăm să abordăm concepții diferite, și expresii matematice variate, pentru a avea o bază mai largă de selecție. Diferite surse argumentează că oamenii de știință tind să aibă ipoteze preferate pentru explicarea proceselor naturale. Ei tind să fie subiectivi atunci când verifică aceste ipoteze, eliminând sau ignorând datele care le infirmă și supraestimând pe cele care le confirmă. Aceasta este un alt fel de a spune că și oamenii de știință sunt supuși aceluiași forțe ca toată omenirea: dacă sunt puși în fața unei explicații care pare satisfăcătoare, comode și logice, sunt motivați să creadă pur și simplu în aceasta. Pentru a ne feri de această continuă sursă de erori și părtinire, se sugerează considerarea totdeauna a multor altor ipoteze de lucru, alternative la posibilele răspunsuri avansate într-un studiu.

După cum am mai scris, Platt (1964, ap. Franklin și col., 2001) scria că, "necesitatea alcătuirii unui sistem ipoteze multiple, alternative, urmate de construirea unui set esențial de experimente care să le verifice, apoi excluderea majorității lor, poartă denumirea de **inferență puternică**, fiind un motor al progresului științei". S-a argumentat că **testarea ipotezelor poate fi benefic înlocuită prin procedura de testare a modelelor**, aceasta fiind abordarea modernă cheie, a procesului de cercetare ipotetic-deductivă.

Din principiul precedent se deduce logic următorul, care sugerează ca întotdeauna **să avem în vedere modele alternative**. Problema selectării unui model valabil este numai parțial legată de complexitate (de exemplu de numărul de variabile independente incluse). La fel de importantă este selectarea a **ce anume se**

include (**care** sunt variabilele independente semnificative în context). Întotdeauna trebuie să avem în vedere că este posibil ca modelului să-i lipsească una sau mai multe trăsături cheie ale sistemului, care explică mai bine comportarea acestuia, decât variabilele considerate până în prezent. Astfel, este benefică testarea presupunerilor și a condițiilor de plecare, și nu numai a prognozelor oferite de model, prin compararea acestora cu datele reale.

Un alt aspect, care nu trebuie uitat, este că modelele statistice au un domeniu de aplicabilitate limitat, și sunt greu sau imposibil de utilizat la alte sisteme sau procese, respectiv în alte condiții. Este ușor să greșim atunci când preluăm un model prefabricat, din literatură sau de pe internet, și să încercăm să-l folosim - de obicei pe date mai puține și variabile mai mult sau mai puțin diferite - într-un alt scop, experiment, habitat și alte condiții ale mediului.

**Confuziile dintre modelele statistice și cele teoretice constituie o sursă inepuizabilă de erori.** Adesea modelul statistic este excelent, explică și prognozează un anumit fenomen cu dezvoltare clară spațio-temporală, după care autorii cad în greșeala considerării acestuia drept model teoretic (deci transpozabil și generalizabil). Însă nu trebuie subestimată posibilitatea dezvoltării judicioase a legăturilor dintre cele două categorii de modele: adesea un model statistic bun ne poate conduce la o serie de raționamente și analize, care vor releva explicativ relațiile intime, cauzale, sau mecanismele legăturilor de tip statistic descoperite. Prin urmare, deși nu sunt sinonime (și nu pot fi niciodată), nu înseamnă că cele două abordări nu sunt legate între ele, sau - mai practic - că nu s-ar sprijini reciproc în calitate de unelte eficiente în cadrul aceluiași demers (idem). Pentru a construi un model eficient sunt necesare seturi de date valabile, specificarea mecanismelor și a proceselor, precum și a parametrilor, dintre care unele sau altele pot lipsi în anumite etape ale unui studiu. Nu trebuie să uităm că modelele sunt (printre altele) metode care permit conducerea, organizarea și orientarea cercetărilor viitoare din laborator sau teren. Ele nu se confundă cu adevărul, ci sunt "**minciuna care ne ajută să vedem adevărul**", după cum se exprima metaforic Fagerstrøm (1987, ap. Mangel și col., 2001).