

Metody analýzy modelů

Radek Pelánek

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Fáze modelování

- 1 Formulace problému
- 2 Základní návrh modelu
- 3 Budování modelu
- 4 Verifikace a validace
- 5 **Simulace a analýza**
- 6 Sumarizace výsledků

- klíčová součást modelování
- především otázka praxe – důležitá část projektu
- přednáška – stručné poznámky:
 - analýza citlivosti
 - vizualizace běhu modelu, výsledků simulací

Poznámky k projektům

- analýza modelu je klíčová součást projektu
- raději jednodušší model s dobrou analýzou než komplikovanější model s povrchní analýzou
- častá chyba – „ověřování hypotéz“

Analýzy, hypotézy, předpoklady

- analýza modelu není ověřování hypotéz
- ověřování hypotéz – typicky skrz randomizované experimenty
- modelování
 - „předpoklady“
 - analýzy ukazují důsledky vložených předpokladů, nemohou je „ověřit“
 - můžeme porovnávat různé modely (předpoklady), zejména pokud máme numerická data

(na základě reálného projektu, avšak zjednodušeno)

- téma: supermarkety vs malé obchody
- „hypotéza“: supermarkety vytlačují malé obchody
- model: pravidla upřednostňující supermarkety
- analýzy: supermarkety rostou, malé obchody klesají
- závěr: hypotéza ověřena

sensitivity analysis

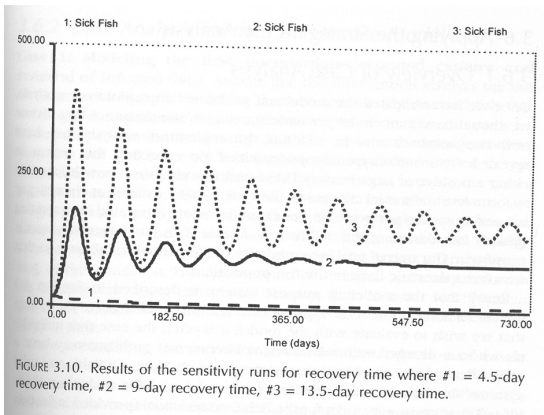
- určení **vlivu parametrů** na chování modelu
- jak moc jednotlivé parametry ovlivňují chování modelu

Analýza citlivosti – důvody

- význam **odhadů**:
 - jak odhady ovlivňují chování modelu?
 - závislost výsledků simulace na správnosti odhadu
- jak **ovlivňovat** chování systému:
 - parametry s vysokým a nízkým vlivem na chování modelu („leverage points“)
 - návrh „politik“ pro změnu chování systému

Příklad

- tři běhy s různými hodnotami jednoho z parametrů
- parametr s vysokým vlivem: ovlivňuje výrazně velikost sledované zásobárny i její průběh (přítomnost/absence oscilací)

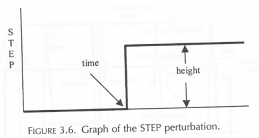
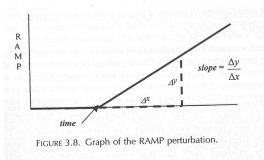
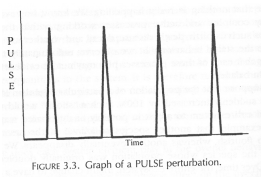


Automatizovaná analýza citlivosti

- simulační nástroje obsahují podporu pro automatizované spouštění více běhů
 - Stella – SensiSpec
 - NetLogo – Behavior Space
- výsledky nejlépe zpracovat v statistickém softwaru (např. R project)

Změny hodnot parametrů

- jak se chová model za dynamicky se měnících podmínek?
- cíle pokusů:
 - zkoumat vliv různých „politik“ na chování systémů (policy analysis)
 - zkoumat robustnost modelu
- příklady experimentů (změna určitého parametru během simulace): Puls, Ramp, Step



Příklady: lovec-kořist

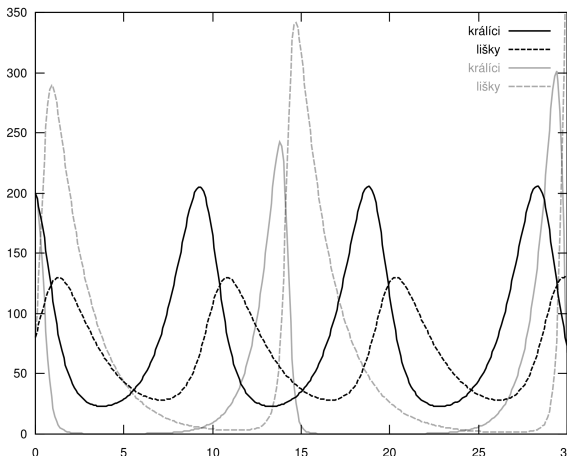
- ABM model králík-tráva (variace na model ze cvičení)
- základní systémový model Lotka-Voltera

$$dL/dt = p_l KL - u_l L$$

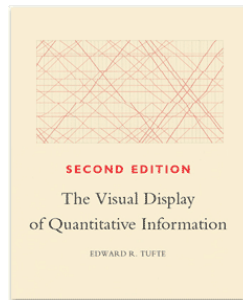
$$dK/dt = p_k K - u_k KL$$

Verifikace modelu

- systémový model (hodnoty parametrů:
 $p_l = 0,006$, $u_l = 0,5$, $p_k = 1$, $u_k = 0,015$)
- černé křivky – numerický výpočet s $\Delta t = 0,001$
- šedé křivky – numerický výpočet s $\Delta t = 0,05$



- doporučená literatura:
E. R. Tufte: The Visual Display of Quantitative Information
- kurzy
 - PV251 Vizualizace
 - PA055 Vizualizace komplexních dat

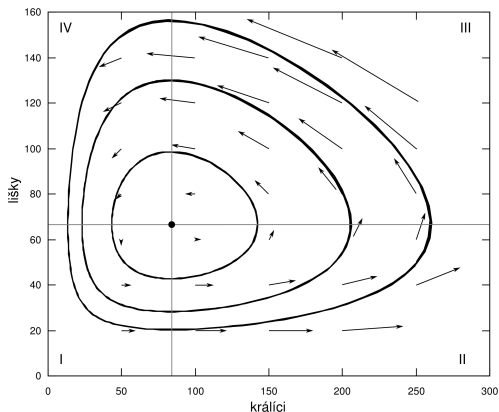


- základní způsob zobrazení dynamického chování
- **přemýšlet**, nepoužívat jen automaticky „default“, např.
 - sumární proměnné – způsob počítání u ABM
 - více proměnných současně (které?)
- důležité věnovat pozornost:
 - délka zobrazovaného intervalu
 - osa y, zvláště při více proměnných

Stavový prostor

- model: vektor délky n
- stavový prostor: n rozměrný prostor
- přímočaré zobrazení pro $n = 2$
- pro vícerozměrné používáme vhodný průmět (vybrané sumární proměnné)

Stavový prostor: systémový model



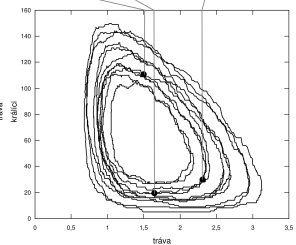
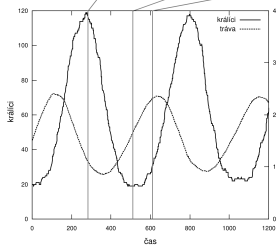
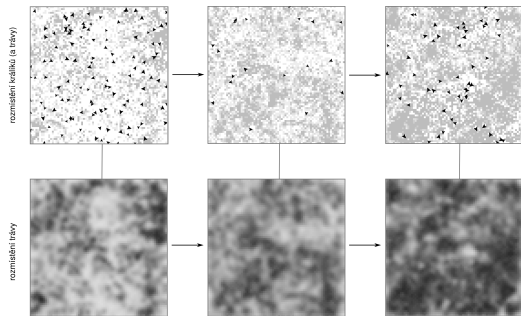
$$dL/dt = p_l KL - u_l L$$

$$dK/dt = p_k K - u_k KL$$

Vizualizace modelu

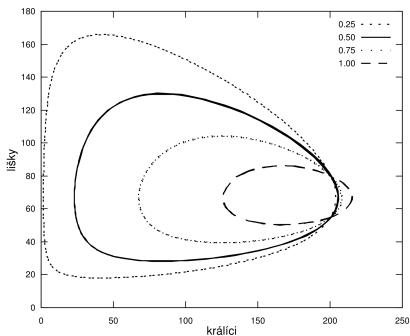
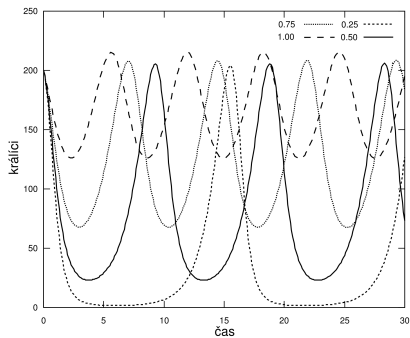
U ABM modelů důležitá vizualizace: ladění, marketing, pedagogické účely, ...

- sledování jednotlivce, zobrazení z pohledu jednotlivce
- numerické informace – velikost objektů, odstín barvy (šedi)
- nedávat všechny informace do jedné vizualizace, nabídnout více různých vizualizací

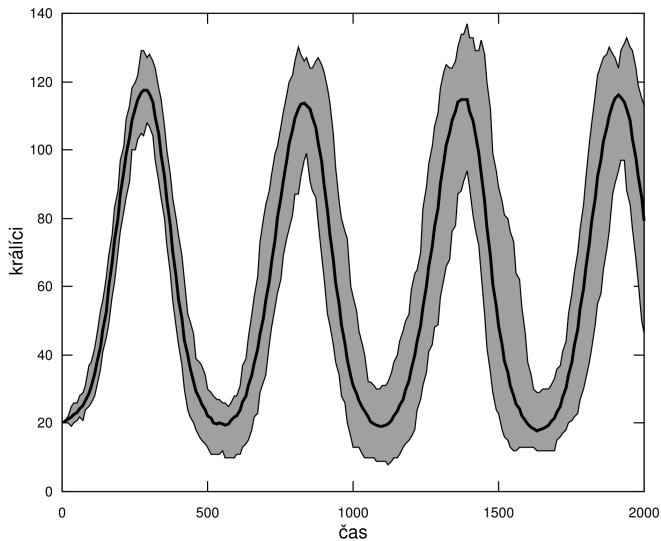


Analýza citlivosti: systémový model

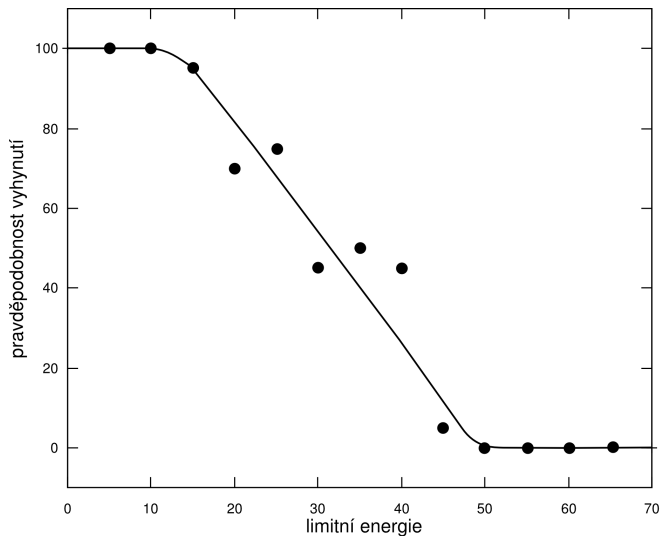
systémový model: změna hodnoty parametru u_1



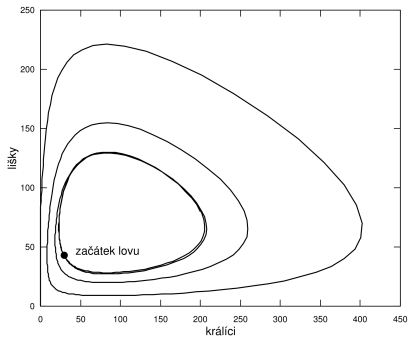
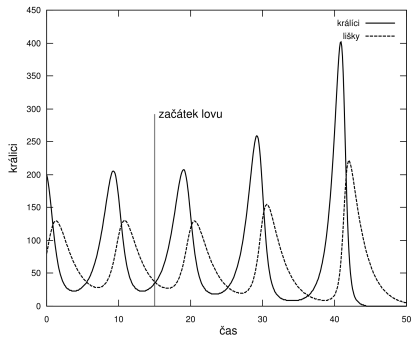
Analýza citlivosti: náhoda u ABM modelu



Analýza citlivosti: závislost na parametru



Příklad: step experiment – skoková změna (lov)



silně doporučený, interaktivní zdroj:

<http://worrydream.com/LadderOfAbstraction/>

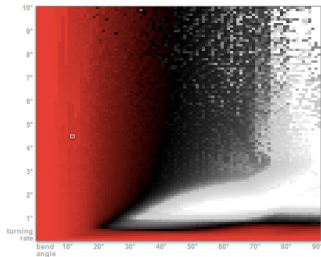
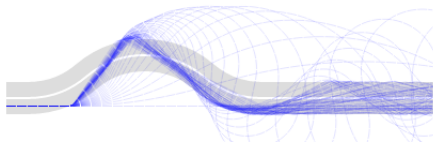
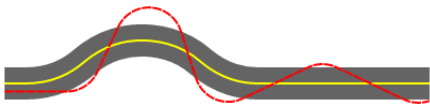
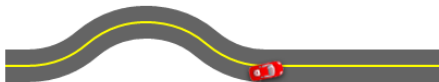
At each step:



Move forward 1 pixel.

If left of the road, turn right by 2° .

If right of the road, turn left by 2° .



Stupně abstrakce

- konkrétní simulace
 - vývoj, ladění
 - pochopení modelu, základní intuice o chování
- abstrakce (znázornění více běhů, abstrahování času nebo některých proměnných)
 - hlubší vhled
 - poučení pro zásahy

Evaluace modelů

- referenční data \Rightarrow vyhodnocení shody modelu a dat
- příklady:
 - Sudoku – obtížnost úloh
 - výukové systémy – odpovědi studentů
 - klima, cena ropy – historická data o vývoji
- souvislosti: strojové učení, přeučení, trénovací/testovací data, ...

Měření shody modelu a dat

data	10	15	13	17	20	25	...
model	8	12	15	17	22	21	...

Jak vyjádřit „kvalitu“ modelu (predikcí) jedním číslem?

Měření shody modelu a dat

data	10	15	13	17	20	25	...
model	8	12	15	17	22	21	...

Jak vyjádřit „kvalitu“ modelu (predikcí) jedním číslem?

- mean absolute error, root mean square error
- correlation (Pearson, Spearman), R^2
- precision, recall, F1
- receiver operating characteristic, area under curve

komplikované zvláště pokud model predikuje pravděpodobnost

- důkladná analýza modelu tvoří klíčovou součást modelování
- není zřejmé jak (co vše) udělat, ani když máme konkrétní data
- časová osa, stavový prostor, vizualizace (ABM), analýza citlivosti, experimenty s modelem
- i s velmi jednoduchým modelem lze provádět zajímavé analýzy
- projekt – důraz právě na analýzy