

**České vysoké učení technické v Praze
Fakulta dopravní**

**Czech Technical University in Prague
Faculty of Transportation Sciences**

Doc. Ing. Emil Pelikán, CSc.

**Predikce procesů v komplexních systémech
Processes Forecasting in Complex Systems**

Summary

The lecture is devoted to the techniques for the models developing suitable for processes forecasting in complex systems. Because of the high sensitivity of the processes to the initial conditions and, consequently, due to our limited possibilities to forecast the processes for the long-term horizon, the attention is focused on the techniques leading to practical applications of the short term prediction models. The aim of this lecture is to bring attention to possible difficulties which designers of the predicting models and their users meet and which have to be solved during the prediction model developing, validation, testing, and applications. The presented overview is not complete, it only reflects the author's experience with developing of the prediction models for practical tasks solving in banking, meteorology, air pollution and energy sector.

The lecture will be completed by an illustration of an example of the global solar radiation prediction which forms an important input for the electrical energy production forecast from renewable sources. The global solar radiation forecasting is based on numerical weather prediction models. The time-lagged ensemble technique for uncertainty quantification is demonstrated on a simple example.

Souhrn

Přednáška je věnována technikám návrhu modelů pro predikci procesů v komplexních systémech. Vzhledem k omezené možnosti predikovat procesy na dlouhodobý horizont z důvodů jejich citlivosti na počáteční podmínky, je pozornost soustředěna na techniky vedoucí k praktické použitelnosti predikčních modelů s krátkodobým předpovědním horizontem. Cílem této přednášky je upozornit na řadu úskalí, se kterými se návrháři i uživatelé předpovědních modelů setkávají. Přehled technik není úplný, vychází pouze ze zkušeností autora získaných při návrhu predikčních modelů pro řešení praktických úloh z oblasti finančnictví, meteorologie a životního prostředí a energetiky.

Přednáška je doplněna příkladem předpovědi globálního slunečního záření, která vstupuje jako významná meteorologická veličina do predikčních modelů pro předpověď výroby elektrické energie z obnovitelných zdrojů. Předpověď globálního slunečního záření je založena na využití numerických modelů počasí. Na jednoduchém příkladu je pak ukázána možnost využití techniky časově zpožděných ensemblů pro kvantifikaci nejistoty předpovědi.

Klíčová slova: predikce, komplexní systémy, asimilace dat, multimodelové predikce, ensemblové predikce, úsudkové předpovídání, globální sluneční záření

Keywords: forecasting, complex systems, data assimilation, multimodel forecasting, ensemble forecasting, judgmental forecasting, global solar radiation

České vysoké učení technické v Praze
Název: Predikce procesů v komplexních systémech
Autor: Doc. Ing. Emil Pelikán, CSc.
Počet stran:
Náklad:

© Emil Pelikán, 2012
ISBN

Obsah

| | | |
|-----------|---|-----------|
| 1 | Úvod | 6 |
| 2 | Obtížnost předpovědi | 7 |
| 3 | Obecný predikční model | 11 |
| 4 | Technika asimilace dat | 12 |
| 5 | Optimální predikční model | 14 |
| 5.1 | Validace modelu | 16 |
| 5.2 | Regularizace | 16 |
| 5.3 | Testování modelu | 17 |
| 6 | Multimodelová predikce | 20 |
| 7 | Ensemblevé předpovědi | 21 |
| 8 | Úsudkové (expertní) předpovídání | 23 |
| 9 | Experiment s předpovědí globálního slunečního záření | 25 |
| 10 | Závěrečná poznámka | 28 |
| 11 | Doc. Ing. Emil Pelikán, CSc. | 31 |

1 Úvod

Studiu procesů probíhajících v komplexních systémech je věnováno v posledních dvou až třech dekadách veliké úsilí jak po teoretické stránce, tak i po stránce aplikované. Komplexní systémy představují relativně novou a široce interdisciplinární oblast jdoucí přes fyzikální, technické, environmentální až po společensko-ekonomické vědy. Tato oblast se poměrně značným tempem rozvíjí zejména v souvislosti s pokroky v oblastech nelineárního modelování, statistické fyziky a počítačových věd. Neexistuje přesná definice toho, co se myslí pod pojmem komplexní systém, ale většina odborníků chápe pod komplexním systémem takový systém, který je tvořen mnoha vzájemně interagujícími složkami (částmi), který vykazuje emergentní chování, tj. kolektivní chování vykazuje rysy, které se nedají jednoduše odvodit anebo vysvětlit z individuálního chování jednotlivých částí. Klasické příklady komplexních systémů můžeme nalézt v ekonomii (modelování vývoje finančních trhů), ekologii (vlivy znečištění životního prostředí), klimatologii (modely globálního oteplování), meteorologii (předpověď počasí), biologii (modely šíření nakažlivých chorob), dopravě (modely dopravních toků), sociologii (populační vývoj) a v řadě dalších oblastí.

Studium chování komplexních systémů je obtížné a pochopení v nich probíhajících procesů je mnohdy velice těžké. To je dáno mimo jiné i složitostí vazeb a mnohdy i nejasností vztahů příčina – následek (viz studium chování makroekonomických veličin jako jsou HDP, nezaměstnanost, inflace apod.). Dalším důvodem je interdisciplinarita, kdy pro pochopení zákonitostí je potřeba znalost reálií z několika vědních oblastí (např. ze společenských věd, informatiky, matematické statistiky, fyziky apod.). I přes rozdílnost různých vědeckých disciplin se ukazuje, že komplexní systémy vykazují společné universální vlastnosti, kterými jsou např. nelineární chování, učící schopnosti apod.

Modelování procesů v komplexních systémech může být rozděleno na dva základní přístupy. První zahrnuje tvorbu a studium základních matematických (a mnohdy zjednodušených, nikoliv však jednoduchých) modelů, pomocí kterých je možné abstrahovat nejdůležitější kvalitativní rysy chování komplexního systému. Zde jsou využívány poznatky z teorie dynamických systémů, sítí, evolučního počítání, numerických metod apod.

Druhým přístupem je vytváření modelů chování z naměřených dat s využitím metod matematické statistiky, statistického učení, dolování znalostí z dat (data mining) apod. Oba přístupy pochopitelně mohou být kombinovány.

Znalost modelů chování nám pak umožní provádět počítačové simulace, vytvářet scénáře budoucího vývoje, predikovat chování na krátký či střednědobý predikční horizont a využívat předpovědi či získané scénáře pro řízení, plánování a rozhodování.

V této přednášce je pozornost soustředěna na techniky, které jsou důležité pro návrh kvalitních modelů pro předpověď chování komplexních systémů, a to zejména na krátkodobý horizont. Přehled zdaleka není úplný, a to jednak z důvodů omezeného rozsahu přednášky a jednak i z důvodu stále intenzivního vývoje nových a zdokonalování stávajících technik a postupů. Výběr prezentovaných technik byl do značné míry dán zkušeností autora při reálném návrhu predikčních modelů a jejich praktického nasazení v oblasti bankovníctví, meteorologie a kvality ovzduší a v energetice.

2 Obtížnost předpovědi

V minulosti bylo vědecké myšlení ovlivněno Newtonovým paradigmatem, kde se svět dá redukovat na několik základních elementárních principů, umožňující jednoduše modelovat a tedy i predikovat chování jednoduchými postupy. Avšak studium chování v mikroskopickém i makroskopickém světě i dalších jevů nám ukázalo, že svět je složitý. Složitost (komplexita) přírodních systémů je přirozenou součástí světa, který nás obklopuje [13]. Můžeme predikovat pohyb planet na mnoho tisíc let dopředu, avšak neumíme spolehlivě předpovědět vývoj HDP na příští rok, počasí na několik dní dopředu, vývoj cen akcií a směnných kurzů na několik hodin či minut dopředu.

Nemožnost přesně (či spíše s dostatečnou a prakticky využitelnou přesností) předpovědět chování komplexních systémů je dána především nekompletní informací o systému, velkým množstvím parametrů a naší nedostatečnou schopností je stanovit, velkým množstvím proměnných a nemožností je změřit a v neposlední řadě i neurčitostí obsaženou v provedených měření (chybou měření).

Základním problémem predikce procesů v komplexních systémech je využití apriorních informací obsažených v dílčím (dostupném) pozorování a známých (přírodních, technických, sociologických a jiných) zákonitostí. Tyto zákonitosti mohou být vyjádřeny jako funkcionální vztah mezi různými uvažovanými veličinami a úlohou je nalézt hodnoty parametrů tak, aby modelové predikce co nejvíce korespondovaly s naměřenými veličinami.

Vyjdeme-li z deterministického popisu jevů, která vychází z teorie dynamických systémů, studované procesy jsou definované jako vektory

časově závislých proměnných (stavů)

$$\mathbf{x}_t = (x_1(t), x_2(t), \dots, x_n(t)), \quad (1)$$

kde dynamika časového vývoje stavů je popsána rovnicí

$$\mathbf{x}_t = \mathbf{F}^t(\mathbf{x}_0, \boldsymbol{\lambda}) \quad (2)$$

a kde \mathbf{x}_t je stav systému v čase t , \mathbf{x}_0 je počáteční stav systému, \mathbf{F}^t je hladká funkce popisující vývoj stavového vektoru v čase a $\boldsymbol{\lambda}$ je vektor (řídících) parametrů. Stav systému mohou záviset jak na čase, tak i na dalších proměnných (prostorová závislost). Podobně i vektor řídících parametrů může záviset na čase případně i na jiných proměnných. Pro jednoduchost však obvykle předpokládáme, že vektor parametrů $\boldsymbol{\lambda}$ nezávisí na čase a že dynamika procesu (2) může být též vyjádřena systémem parciálních diferenciálních rovnic

$$\frac{\partial x_i(t)}{\partial t} = f_i(\mathbf{x}_t, \boldsymbol{\lambda}) \quad (3)$$

kde funkce f_i také na čase nezávisí. Známe-li funkce f_i a dokážeme-li nějakým způsobem určit hodnotu řídících parametrů $\boldsymbol{\lambda}$, je otázkou, proč by měla predikce časového vývoje takového deterministického systému představovat nějaký problém.

K tomuto existuje v literatuře, viz např. [2], [7], názorný příklad velmi jednoduchého (tzv. logistického) modelu popisujícího populační vývoj, který je dán rovnicí

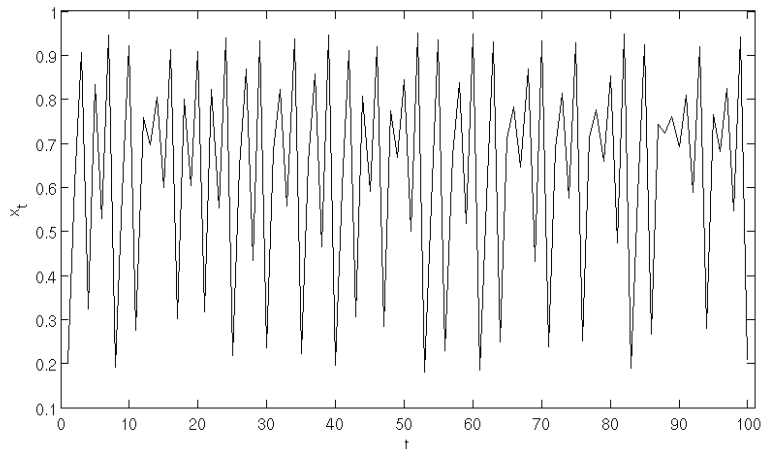
$$\frac{dx_t}{dt} = \alpha x_t \left(1 - \frac{x_t}{K}\right), \quad (4)$$

kde x_t je velikost populace v daném omezeném prostoru, α je parametr procesu a K je saturační koeficient. V logistickém modelu je zakomponována apriorní informace o tom, že změna velikosti populace je úměrná velikosti populace a dále, že v omezeném prostoru existuje saturační mez.

Pro diskrétní časový krok lze uvažovat o rovnici

$$x_{t+1} = \alpha x_t \left(1 - \frac{x_t}{K}\right). \quad (5)$$

Ačkoliv je logistický model jednoduchý, chování systému popsaného rovnicí (5) vůbec jednoduché není. Na obrázku 1 je uveden průběh časové řady generované logistickým modelem pro prvních 100 hodnot s počáteční podmínkou $x_0 = 0,2$ a s parametry $K = 1$ a $\alpha = 3,8$.



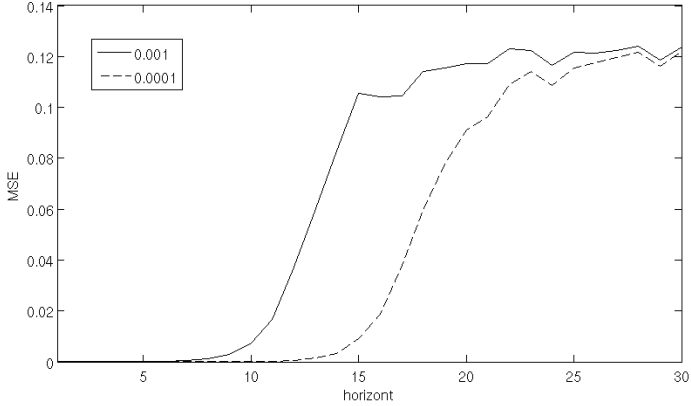
Obrázek 1: Průběh časové řady (prvních 100 hodnot) generované logistickou funkcí s počáteční podmínkou $x_0 = 0,2$

Dá se ukázat [2], že pro širokou třídu hodnot parametrů α vykazuje systém (5) chaotické chování (Ljapunovův exponent systému je kladný). To mimo jiné znamená, že chování systému bude velmi citlivé na hodnoty počátečních podmínek x_0 . Tento fakt je demonstrován na obrázku 2, kde je uveden průběh střední kvadratické chyby předpovědi chování systému (5) pro různé počáteční podmínky v závislosti na předpovědním horizontu

$$\text{MSE}(m) = \frac{1}{N} \sum_{i=1}^N [F^m(x_0) - F^m(x_0 + \varepsilon_i)]^2, \quad (6)$$

kde $F^m(x_0)$ je vývoj systému (5) z počáteční podmínky x_0 na m kroků dopředu a ε_i jsou náhodně vybrané hodnoty z rovnoměrného rozdělení na intervalu $(-0,0001; +0,0001)$, respektive $(-0,001; +0,001)$. Pro výpočet bylo použito 10 000 realizací náhodné chyby pro každou počáteční podmínku x_0 , přičemž počáteční podmínky nabývaly různých hodnot od 0,005 do 0,995 a výsledná střední kvadratická chyba byla zprůměrována.

Z obrázku 2 je názorně vidět, jak chyba předpovědi exponenciálně narůstá s rostoucím předpovědním krokem (až saturačnímu limitu). Provedený experiment dokládá obtížnost předpověditelnosti chování chaotických komplexních systémů na delší předpovědní horizont, pokud neznáme počáteční stav dostatečně přesně, a to i za předpokladu,



Obrázek 2: Průběh střední kvadratické chyby MSE v závislosti na predikčním horizontu m pro logistickou funkci $x_{t+1} = 3,8x_t(1 - x_t)$ s náhodně zašuměnými počátečními podmínkami x_0

že dynamiku systému dokonale známe, a to včetně hodnot zahrnutých parametrů. Ostatně vliv nepřesnosti v odhadu parametru může způsobit také exponenciální nárůst predikční chyby, jak je demonstrováno na obrázku 3.

Problém nemožnosti změřit stav systému s dostatečnou přesností se v některých případech komplikuje tím, že z technických či principiálních důvodů nemůžeme měřit většinu stavových proměnných a měření je prováděno pouze na velmi omezené části stavových proměnných. Z teorie dynamických systémů vyplývá, že dynamiku systému lze do jisté míry zkoumat pouze z dynamiky vývoje jediného stavu. Pokud uvažujeme deterministický systém dimenze n se stavu

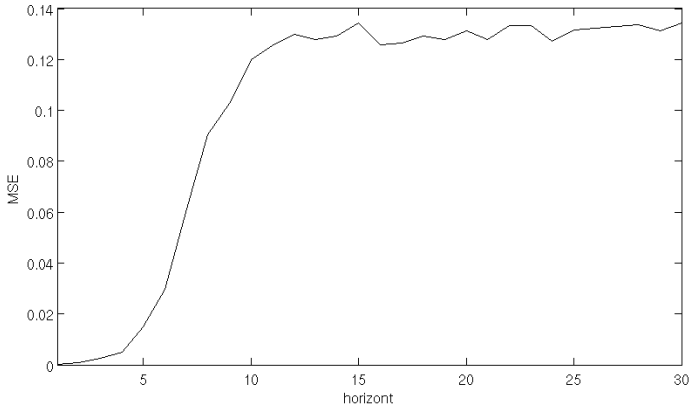
$$\mathbf{x}_t = (x_1(t), x_2(t), \dots, x_n(t)), \quad (7)$$

z Takensova teorému [20] vyplývá, že dynamiku tohoto systému (atraktor) lze zrekonstruovat z vektorů časově posunutých (o krok T) stavů jedné proměnné s vnořenou dimenzí d_E

$$\mathbf{x}_{t,1} = (x_1(t), x_1(t+T), \dots, x_1(t+[d_E-1]T)), \quad (8)$$

kde $d_E > 2n + 1$. Pochopitelně pokud je n hodně veliké, je vnořená dimenze ještě větší. Přesto v řadě aplikací může být zkoumána dynamika celého systému pomocí vnořených vektorů s mnohem nižší dimenzí [7].

Předpovídání procesů v systémech s chaotickým chováním je tedy obtížné, ale ne nemožné. Je jen potřeba respektovat některá omezení,



Obrázek 3: Průběh střední kvadratické chyby MSE pro logistickou funkci $x_{t+1} = \lambda x_t(1 - x_t)$ s různými hodnotami parametru λ generovaného náhodně z intervalu $(3,7;3,9)$ v závislosti na predikčním horizontu m .

a to zejména velkou citlivost na počáteční podmínky a z ní plynoucí nemožnost dosažení nízké predikční chyby na delší časový horizont, dále nutnosti použití obecně nelineárních modelů, složitějších postupů a technik vedoucích k úspěšné predikci.

3 Obecný predikční model

Jak bylo naznačeno v předcházející části, základním problémem předpovídání chování složitých systémů je nemožnost změřit stavy systému s dostatečnou přesností. Běžným cílem návrhu predikčních metod je predikovat časový vývoj stavů komplexního systému obvykle na krátkodobý horizont (viz problémy s předpovědí systémů s chaotickým chováním na delší časový horizont popsané v předcházející části), a to za reálného předpokladu, že nemáme anebo nemůžeme využít kompletní informaci o dynamice systému, parametry musíme odhadovat z naměřených dat a navíc dostupná měření jsou zatížena měřicí chybou. Kromě toho se často setkáváme se situací, kdy nejsme schopni některé stavy systému (a to mnohdy i výraznou většinu stavů) vůbec změřit resp. pozorovat (viz následující kapitola o asimilaci dat). Omezíme-li se na diskrétní systémy a na predikční horizont jeden krok dopředu, lze úlohu predikce zjednodušit na nalezení predikční funkce (predikčního

modelu) F^f

$$\mathbf{x}_t^f = F^f(\mathbf{x}_{t-}, \boldsymbol{\theta}) \quad (9)$$

kde \mathbf{x}_{t-} je historie (naměřených) stavů

$$\mathbf{x}_{t-} = (\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_0) \quad (10)$$

a $\boldsymbol{\theta}$ je vektor parametrů. Potřebujeme-li predikovat na k kroků dopředu ($k > 1$), lze predikční model F^f využít v následujícím tvaru:

$$\mathbf{x}_{t+k-1}^f = F^f(\mathbf{x}_{t+k-2}^f, \mathbf{x}_{t+k-3}^f, \dots, \mathbf{x}_t^f, \mathbf{x}_{t-}, \boldsymbol{\theta}). \quad (11)$$

V reálné praxi se často omezujeme na predikční modely, které nevyužívají celou naměřenou historii stavů, ale pouze na určitou část aktuální historie, tedy například

$$\mathbf{x}_t^f = F^f(\mathbf{x}_{t-,h}, \boldsymbol{\theta}), \quad (12)$$

kde

$$\mathbf{x}_{t-,h} = (\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_{t-h}) \quad (13)$$

Návrh konkrétní podoby predikčního modelu F^f může být založen na mnoha principech vycházejících například z Box-Jenkinsovy metodologie, principů umělých neuronových sítí, fuzzy přístupů, Bayesovských a SVM (support vector machines) technik a z řady dalších. Toto rozsáhlé a zajímavé téma však není cílem této přednášky a proto odkazujeme zájemce na bohatou odbornou literaturu, např. [1], [3], [11], [16], [19].

4 Technika asimilace dat

Ať už bude obecný predikční model popsán v předcházející části v jakémkoliv tvaru, je potřeba zajistit dostupnost vstupních hodnot, tedy stavů \mathbf{x}_{t-} potřebných pro výpočet predikce \mathbf{x}_t^f . Jak už bylo řečeno výše, v některých systémech z technických či principiálních důvodů nemůžeme měřit většinu stavových proměnných a měření je prováděno pouze na velmi omezené části stavových proměnných. Pokud potřebujeme zrekonstruovat všechny stavy systému z dostupných měření (která pochopitelně obsahují šum), lze použít techniku asimilace dat. Ta je založena na modifikaci prvotních (apriorních) odhadů stavů podle chyby observačního operátoru a dostupných měření (viz např. [9]). Předpokládáme-li, že prvotní odhad stavů v systému je \mathbf{x}_t^p , \mathbf{H} je observační operátor transformující (popř. interpolující) stavy systému \mathbf{x}_t^p na množinu

pozorovaných stavů \mathbf{x}_t^o , pak nově zrekonstruovaný stav systému \mathbf{x}_t^{new} s využitím měření \mathbf{x}_t^o je dán vztahem

$$\mathbf{x}_t^{new} = \mathbf{x}_t^p + \mathbf{W} [\mathbf{x}_t^o - \mathbf{H}(\mathbf{x}_t^p)], \quad (14)$$

kde matice \mathbf{W} je váhová matice, která je obvykle určena na základě odhadů kovariancí statistických chyb pozorování a stavů systému. Pro prvotní odhad stavů systému může být použita varianta predikčního modelu (9), tedy pokládáme

$$\mathbf{x}_t^p = \mathbf{x}_t^f. \quad (15)$$

Existuje celá řada metod asimilace dat založených na vztahu (14), jako jsou metody založené na technice optimální interpolace (OI), 3D-Var, 4D-var [4], [9].

Jako příklad zde uvedeme jednu z nejužívanějších metod, a to metodu s využitím techniky Kalmanových filtrů. Předpokládáme, že stavy systému jsou náhodné \mathbf{x}_t , a že příslušný náhodný proces má Markovskou vlastnost, tj. je splněna rovnost

$$p(\mathbf{x}_t | \mathbf{x}_{t-h}) = p(\mathbf{x}_t | \mathbf{x}_{t-1}). \quad (16)$$

Předpokládejme dále, že rekonstrukce stavů \mathbf{x}_{t-1}^{new} je nevychýleným odhadem, tedy že platí

$$\mathbf{x}_{t-1}^{new} = \mathbf{x}_{t-1} + \boldsymbol{\eta}_{t-1}, \quad (17)$$

kde střední hodnota náhodných chyb $E(\boldsymbol{\eta}_{t-1}) = \mathbf{0}$, kde E je operátor střední hodnoty. Předpokládáme dále, že predikční model použitý pro předpověď prvotního stavu systému i observační operátor \mathbf{H} jsou lineární, tedy

$$\mathbf{x}_t^p = \mathbf{M}_t \mathbf{x}_{t-1}^{new} \quad (18)$$

a

$$\mathbf{H}(\mathbf{x}_t^p) = \mathbf{H}_t \mathbf{x}_t^p, \quad (19)$$

kde \mathbf{M}_t , \mathbf{H}_t jsou matice parametrů. Označme dále

$$\boldsymbol{\varepsilon}_t = \mathbf{x}_t^o - \mathbf{H}_t \mathbf{x}_t, \quad (20)$$

$$\boldsymbol{\nu}_t = \mathbf{x}_t - \mathbf{M}_t \mathbf{x}_{t-1}. \quad (21)$$

Kovarianční matice chyb označíme

$$\text{cov}(\boldsymbol{\eta}_{t-1}) = \mathbf{P}_{t-1}^{new}, \quad (22)$$

$$\text{cov}(\boldsymbol{\nu}_t) = \mathbf{Q}_t, \quad (23)$$

$$\text{cov}(\boldsymbol{\varepsilon}_t) = \mathbf{R}_t. \quad (24)$$

Za předpokladu, že ε_t a ν_t jsou nekorelované a mají též nulovou střední hodnotu, lze odvodit následující rekurzivní vztahy

$$\mathbf{x}_t^{new} = \mathbf{x}_t^p + \mathbf{K}_t (\mathbf{x}_t^o - \mathbf{H}_t \mathbf{x}_t^p), \quad (25)$$

$$\mathbf{P}_t^{new} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_t^p, \quad (26)$$

$$\mathbf{P}_t^p = \mathbf{M}_t \mathbf{P}_{t-1}^{new} \mathbf{M}_t^T + \mathbf{Q}_t, \quad (27)$$

kde

$$\mathbf{K}_t = \mathbf{P}_t^p \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t^p \mathbf{H}_t^T + \mathbf{R})^{-1}. \quad (28)$$

Tyto vztahy představují rovnice klasického Kalmanova filtru, přičemž matice \mathbf{K}_t je označována jako Kalmanovo zesílení (Kalman gain).

Z hlediska návrhu modelů pro předpověď procesů v komplexních systémech je dobré si všimnout, že predikční model (18) je lineární, avšak matice parametrů může záviset na čase (tzv. adaptivní model).

Jsou-li všechna uvažovaná rozdělení normální, je výsledné rozdělení

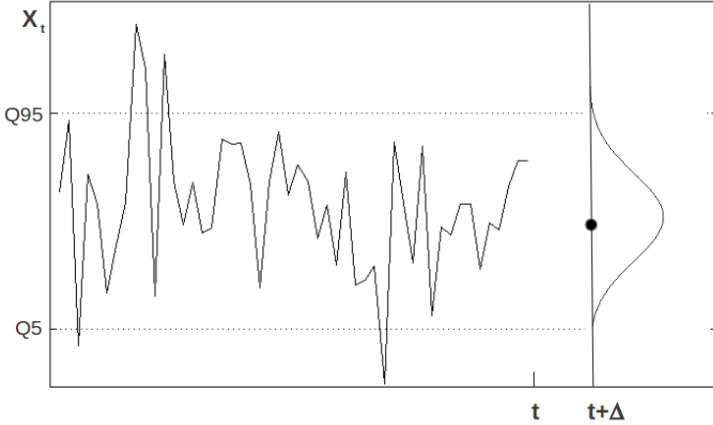
$$p(\mathbf{x}_t^{new} | \mathbf{x}_t^o) \quad (29)$$

také normální se střední hodnotou \mathbf{x}_t^{new} a kovarianční maticí \mathbf{P}_t^{new} . Za výše uvedených předpokladů lze tedy relativně snadno předpovídat nejen „střední hodnotu“ nového stavu, ale i celé pravděpodobnostní rozdělení (viz obr. 4) a z něj odvozených hladin spolehlivosti (kvantilová předpověď). Rovnice Kalmanova filtru (18)–(28) umožňují též předpověď měřených stavů \mathbf{x}_t^o a jsou základem techniky tzv. state-space predikčních modelů, viz např. [3], [16].

5 Optimální predikční model

V komplexních systémech obvykle lze pro stejnou úlohu navrhnout celou řadu predikčních modelů, které mohou být využity pro reálnou předpověď. Tak jako v mnoha vědních oblastech nejlepší možné řešení (optimální model) obvykle není možné nalézt a proto je návrh predikčních modelů invenčním procesem. Zde by se hodil citát připisovaný G. Boxovi a W.E. Demingovi: „Všechny modely jsou špatné, některé jsou užitečné.“ [14].

Pokud máme vhodnou kriteriální funkci, která ohodnocuje kvalitu předpovědi našeho modelu danou nesouladem (diskrepancí) mezi predikovanými a skutečnými hodnotami, lze nejlepší model posuzovat podle hodnoty této kriteriální funkce. Základním problémem je, že hodnotu této funkce umíme vyčíst pouze na základě známých skutečností (na základě historických dat). Abychom věřili, že náš model bude nejlepší



Obrázek 4: Bodová, pravděpodobnostní předpověď a předpověď 5% a 95% kvantilů - ilustrativní obrázek.

i budoucnu – a to je to, o co nám především jde – musíme učinit jisté předpoklady o našem systému a o datech, která generuje.

Pokud uvažujeme predikční model ve tvaru $\mathbf{x}_t^f = \mathbf{F}^f(\mathbf{x}_{t-h}, \boldsymbol{\theta})$, je základní úlohou návrh struktury predikční funkce a odhad jejích parametrů. Struktura predikční funkce může být určena apriori (na základě apriorních znalostí o chování systému) anebo z dat např. vhodnou volbou třídy dostatečně universálních funkcí tak, jako je tomu například při použití umělých neuronových sítí. Parametry $\boldsymbol{\theta}$ pak mohou být určeny na základě minimalizace vhodně zvolené kriteriální resp. ztrátové (loss-function) funkce. Pro konkrétní naměřené hodnoty stavových veličin \mathbf{x}_t , $t = 1, 2, \dots, N$ lze uvažovat o ztrátové funkci typu

$$L(\boldsymbol{\theta}) = \|\mathbf{x}_t - \mathbf{F}^f(\mathbf{x}_{t-h}, \boldsymbol{\theta})\|^2. \quad (30)$$

Protože hledáme takový parametr $\boldsymbol{\theta}$, pro který je (30) minimální, můžeme napsat

$$\hat{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} L(\boldsymbol{\theta}). \quad (31)$$

Tato procedura se nazývá minimalizace empirického rizika (empirical risk minimization – ERM). Protože predikční funkce \mathbf{F}^f může být velmi komplikovaná (nelineární apod.), hledání vhodného $\boldsymbol{\theta}$ minimalizujícího $L(\boldsymbol{\theta})$ může být netriviální úloha. Základní problém procedury ERM je však ten, že odhad parametru $\boldsymbol{\theta}$ na základě dostupných dat

\mathbf{x}_t , $t = 1, 2, \dots, N$, zřejmě nebude identický s odhadem provedeným na datech naměřených v následujícím období, a to i za předpokladu, že funkce \mathbf{F}^f se v čase nemění. Může se tedy stát, že hodnota ztrátové funkce $L(\boldsymbol{\theta})$ v následujícím období vzroste tak, že se model stane neúčinným. Zajímá nás tedy nejen hodnota $L(\boldsymbol{\theta})$ na daných datech, která se nazývá trénovací chyba (training error anebo in-sample error), ale také informace, jak se bude vyvíjet ztrátová funkce na datech, která teprve budou získána. Tento fakt můžeme vyjádřit tím, že nás tedy zajímá „očekávaná hodnota“ ztrátové funkce $E[L(\boldsymbol{\theta})]$, kde E je operátor střední hodnoty. Funkce $E[L(\boldsymbol{\theta})]$ se nazývá generalizační chyba (generalization error anebo out-of sample error). Důvodem rozdílnosti mezi trénovací chybou a generalizační chybou je fakt, že trénovací data \mathbf{x}_t , $t = 1, 2, \dots, N$ nejsou reprezentativní pro chování systému, který studujeme (ale jen nějakou ukázkou chování danou výběrem dat získaných s technickými, fyzikálními či jinými omezeními) a dále že dostupná data (měření) nejsou perfektní (obsahují šum).

Problém přetrénování (overfitting) modelu lze do jisté míry odstranit použitím technik validace modelu (out-of sample testing) a regularizací minimalizačního procesu.

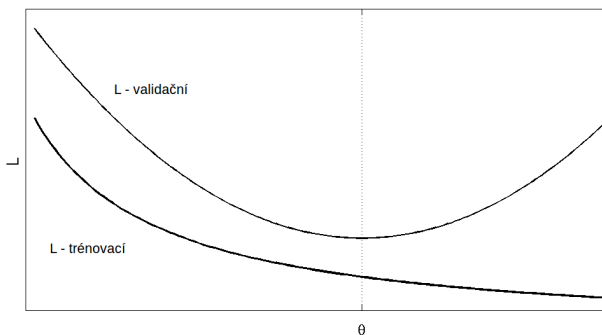
5.1 Validace modelu

Při validaci modelu provádíme odhad generalizační chyby s využitím druhého, pokud možno nezávislého výběrového souboru dat, který nevstupuje do minimalizačního procesu (31). Obvykle tedy rozdělíme dostupná data na trénovací a validační, na kterých vyhodnocujeme hodnoty ztrátové funkce. Rozdělit data na trénovací a validační lze různým způsobem (náhodně, postupným výběrem určité části dat apod.).

Pokud používáme numerické iterační procedury v proceduře ERM (minimalizaci ztrátové funkce), lze sledovat průběh empirické chyby i odhad generalizační chyby v průběhu iteračního procesu. Zatímco empirická chyba obvykle klesá (v závislosti na zvolené numerické proceduře), odhad generalizační chyby ztrátové funkce vypočtené na validačním souboru dat po určité době obrací trend (z klesajícího na rostoucí), jak ilustruje obrázek 5. V tomto bodě je vhodné iterační proces zastavit a použít jako odhad parametru takové $\boldsymbol{\theta}$, které minimalizuje průběh odhadu generalizační chyby (early stopping kritérium).

5.2 Regularizace

Problém minimalizace $L(\boldsymbol{\theta})$ může být a většinou i bývá špatně podmíněn, tzn. při malých změnách ve vstupních datech můžeme dostávat



Obrázek 5: Odhad parametru θ s využitím průběhu ztrátové funkce na trénovacích a validačních datech (early-stopping kritérium).

velké změny v odhadech parametrů θ . Proto se minimalizuje

$$L(\theta) + \lambda d(\theta), \quad (32)$$

kde funkce $d(\theta)$ je regularizační anebo penalizační funkce.

Vhodnou volbou regularizační (penalizační) funkce pak můžeme redukovat komplexitu modelu a tím zmenšovat generalizační chybu modelu. Tento princip je v souladu jak s intuitivním chápáním toho, že komplexní (přeparametrizované) modely mohou se zvýšenou měrou vystihovat vlastnosti výběrových učících dat (včetně šumu v nich obsaženého) a nikoliv obecné vlastnosti procesu, který je potřeba predikovat. Současně je tento princip v souladu s filozofickým pojetím známého principu Occamovy břitvy, který říká, že „není třeba rozmnožovat entity, pokud to není nezbytné“, tedy, že z modelů s přijatelnou predikční výkonností na trénovacích datech vyberme ten, který je nejjednodušší.

Příkladem penalizační funkce je funkce

$$K \frac{C(N)}{N}, \quad (33)$$

kde K je počet parametrů modelu a N je rozsah dat. Pro $C(N) = 2$ pak dostáváme často používané Akaikeho informační kritérium AIC, [16].

5.3 Testování modelu

Pokud je model navržen s využitím trénovacích a validačních dat, doporučuje se jej nezávisle otestovat na testovacím souboru, který neob-

sahuje data z trénovací ani z validační množiny. Kriteriaální (ztrátová) funkce pro testování modelu může mít mnoho podob reflektujících požadavky uživatelů na kvalitu predikčních modelů, jejichž splnění nemusí být triviální. Je dobré si také uvědomit, že v řadě praktických aplikací kriteriaální funkce pro testování modelů nemusí být totožná s kriteriaální funkcí pro návrh modelů v etapě trénování a validace.

Pro jednoduchost zde předpokládejme, že našim cílem je predikce pouze jedné komponenty (stavové veličiny) komplexního systému a že testujeme predikční model na testovacím souboru x_t , $t = 1, 2, \dots, N$.

Střední kvadratická chyba MSE (mean square error) je pak definovaná rovnicí

$$L_{MSE} = \frac{1}{N} \sum_{t=1}^N (x_t - x_t^f)^2, \quad (34)$$

kde x_t^f je předpověď veličiny x_t . Porovnávat kvalitu modelů podle MSE má řadu výhod, neboť minimalizace L_{MSE} (tzv. metoda nejmenších čtverců) se hojně používá pro návrh predikčních modelů. Zobecněním kriteriaální funkce (34) je Minkovského kriteriaální funkce, která je definována vztahem

$$L_M = \frac{1}{N} \sum_{t=1}^N |x_t - x_t^f|^R. \quad (35)$$

Je zřejmé, že pro $R = 2$ je funkce (35) identická s funkcí (34). Pro $R = 1$ dostáváme střední absolutní chybu MAE (mean absolute error).

Abychom mohli vzájemně porovnávat predikční schopnosti modelů pro předpověď veličin s různou střední hodnotou (např. při předpovědi makroekonomických veličin pro různě velké regiony), používá se velmi oblíbená střední procentní chyba MAPE (mean absolute percentage error) definovaná vztahem

$$L_{MAPE} = \frac{100}{N} \sum_{t=1}^N \left| \frac{x_t - x_t^f}{x_t} \right|. \quad (36)$$

Ve jmenovateli L_{MAPE} se vyskytuje absolutní hodnota predikované veličiny, která při velmi malých hodnotách blízkých nule vede k velmi vysokým hodnotám tohoto kritéria, proto je potřeba procentní chyby vyhodnocovat s velkou opatrností.

Vyhodnocení kvality modelu podle velkých odchylek je jiným příkladem kriteriaální funkce vyjádřené vztahem

$$L_{MAX} = \frac{1}{M} \sum_{|x_t - x_t^f| > \varepsilon} |x_t - x_t^f|, \quad (37)$$

kde ε je předem zvolená přípustná hodnota chyby a M je počet predikčních chyb větších (v absolutní hodnotě) než ε .

V odborné literatuře o predikcích [11] se dále vyskytuje kritérium, které porovnává výkonnost modelu s tzv. persistentním modelem

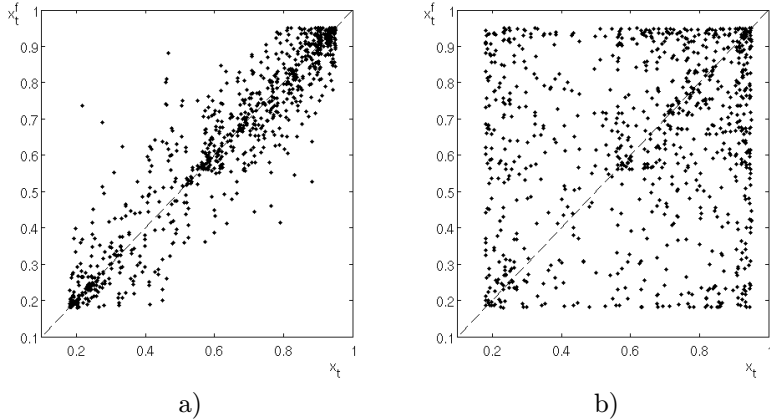
$$x_t^f = x_{t-1}. \quad (38)$$

Ze vztahu (38) vyplývá, že persistentní model predikuje nejbližší budoucnost posledním známým stavem systému. Kriteriaální funkce (tzv. Theilova U -statistika) je pak definována vztahem

$$L_U = \sqrt{\frac{\sum_{t=1}^N \left(\frac{x_{t+1} - x_{t+1}^f}{x_{t+1}} \right)^2}{\sum_{t=1}^N \left(\frac{x_{t+1} - x_t}{x_t} \right)^2}}. \quad (39)$$

Pro modely s lepší výkonností než je persistentní model je pak hodnota $L_U < 1$. Obdobným způsobem pak můžeme porovnávat výkonnost predikčního modelu s modelem typu náhodná procházka apod.

Jako poslední příklad vyhodnocení kvality modelu uvedme použití křížových diagramů, které jsou zobrazeny na grafech v obrázku 6



Obrázek 6: Křížový diagram predikované versus skutečné hodnoty pro logistickou funkci $x_{t+1} = 3,8x_t(1 - x_t)$ s náhodně zašuměnými počátečními podmínkami x_0 pro predikční krok $m = 10$ (a) a $m = 15$ (b).

6 Multimodelová predikce

Složitost předpovídání procesů v komplexních systémech způsobuje to, že lze vytvořit více modelů (založených na stejných anebo odlišných principech). Modely lze testovat na učicích i validačních datech a vybrat z portfolia modelů ten nejlepší. Tvorba modelů je však náročný proces a tak se naskytá otázka, zda nevyužít vlastností i ostatních (již vytvořených) modelů. Na tom je založen princip multimodelových předpovědí, kdy se kombinují predikce z různých modelů, přičemž se často stává, že finální predikce vykazuje lepší výkonnost (menší predikční chybu) než je predikční výkonnost jednotlivých individuálních modelů.

Předpokládáme-li, že multimodelová predikce je tvořena aritmetickým průměrem K předpovědí z K predikčních modelů

$$x_t^{MIX} = \frac{1}{K} \sum_{j=1}^K x_{t,j}^f, \quad (40)$$

pak střední kvadratická chyba kombinované předpovědi je rovna

$$E(x_t^{MIX} - x_t)^2 = E\left(\frac{1}{K} \sum_{j=1}^K x_{t,j}^f - x_t\right)^2 = E\left(\frac{1}{K} \sum_{j=1}^K \varepsilon_{t,j}\right)^2, \quad (41)$$

kde $\varepsilon_{t,j}$ je predikční chyba j -tého predikčního modelu. Za předpokladu, že chyby jednotlivých modelů jsou nekorelované s nulovou střední hodnotou a s identickým rozptylem σ^2 , pak platí

$$E(x_t^{MIX} - x_t)^2 = \frac{1}{K^2} \sum_{j=1}^K E(\varepsilon_{t,j}^2) = \frac{1}{K} \sigma^2, \quad (42)$$

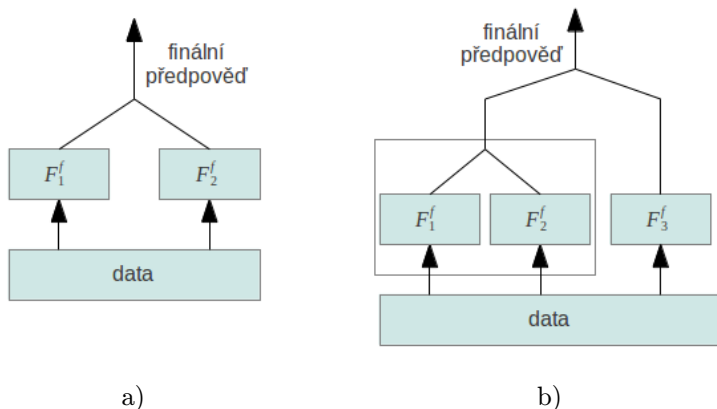
kde E je operátor střední hodnoty. Rozptyl výsledného chybového procesu je tak redukován (za daných předpokladů) faktorem K . V obecném případě jsou chyby jednotlivých modelů korelované, avšak i tak v mnoha případech lze multimodelovou předpověď predikční chyby redukovat.

Pochopitelně lze vzorec pro kombinování (40) zobecnit tak, že finální predikce je váženou kombinací jednotlivých predikcí

$$x_t^{MIX} = \sum_{j=1}^K w_j x_{t,j}^f, \quad (43)$$

kde váhy jsou navrženy tak, aby minimalizovaly hodnotu příslušné ztrátové funkce.

Výše uvedený princip lze použít k aktivnímu návrhu individuálních modelů vhodných pro kombinovanou předpověď například tak, že k danému portfolio modelů je navržen další, přičemž požadujeme, aby jeho predikční chyba byla „dekorelována“ s chybou stávajícího portfolio [15]. Schématicky je tento postup zobrazen na obrázku 7. Lze také jak váhy, tak i parametry individuálních modelů odhadovat současně (viz tzv. mixture of experts models [1]).

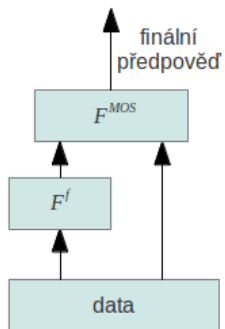


Obrázek 7: Schéma klasické kombinované předpovědi dvěma predikčními modely (a) a přidání nového modelu ke stávajícímu portfolio (b).

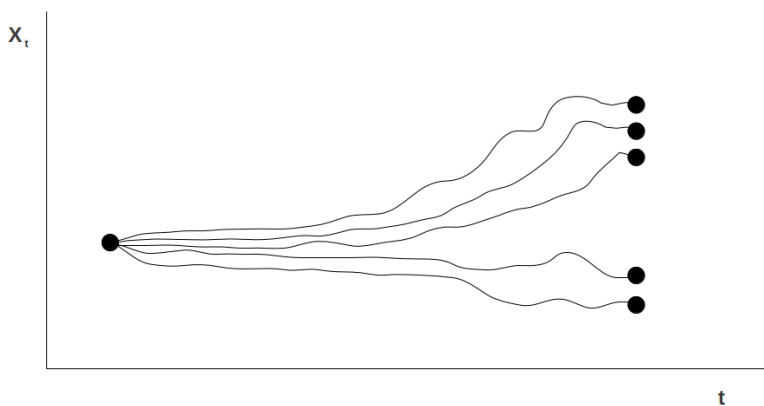
Do kategorie multimodelového předpovídání můžeme zahrnout i techniku zvanou MOS (model output statistics), která je založená na statistickém zpracování (postprocessing) výstupů z fyzikálních (biologických, chemických aj.) modelů. Predikce z těchto modelů (často velmi sofistikovaných) mohou být vychýlené, predikční chyby jsou korelované apod. Cílem techniky MOS je navrhnout statistický model vylepšující statistické charakteristiky původních modelů. Schématicky je použití techniky MOS ukázáno na obr. 8.

7 Ensemblevé předpovědi

Jak bylo diskutováno v kapitole 2, vlastností komplexních systémů je velká citlivost dalšího vývoje na počátečních podmínkách. Podstatou ensemblevého předpovídání je získání více individuálních předpovědí (předpokládaných trajektorií) vhodnou perturbací počátečního stavu systému. Schématicky je tento postup ukázán na obrázku 9.



Obrázek 8: Statistické vylepšení predikčních vlastností predikčního modelu F^f technikou MOS (model output statistics).



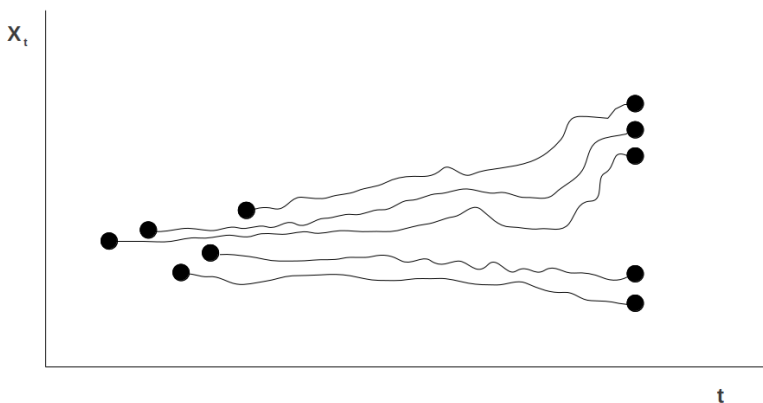
Obrázek 9: Získání ensemblových předpovědí perturbací počátečního stavu komplexního systému.

Z těchto předpovědí se dá získat jednak finální předpověď (například jako vážená kombinace individuálních předpovědí obdobně jako v případě multimodelových předpovědí popsanych v předcházející části), jednak informace o nejistotě daných předpovědí. Získané ensemblové předpovědi by ideálně měly být náhodným výběrem z pravděpodobnostního rozdělení, které odpovídá vývoji daného systému. Nejlepším řešením by bylo popsat časový vývoj zkoumaného systému pomocí rovnic popisující vývoj hustoty pravděpodobnosti $p(\mathbf{x}_t)$ tak, jak se to zkoumá v teorii stochastických procesů a jak bylo naznačeno v kapitole 4 a předpověď poskytovat ve formě predikované hustoty pravděpodobnosti. V reálných situacích je však často obtížné splnit všechny nezbytné předpoklady a spolehlivě odhadnout teoretickou hustotu pravděpodobnosti. Ensemblové předpovědi tak představují užitečný krok od „bodových předpovědí“ k plně pravděpodobnostním predikcím. Ostatně z praktického hlediska (podle autorových zkušeností) není jednoduché běžnému uživateli přístupnou formou sdělit výstupy z pravděpodobnostních předpovědních modelů.

Jinou metodou je konstrukce tzv. časově zpožděného ensemblu (time-lagged ensemble), kde ensemble individuálních předpovědí pro daný čas $t+k$ je vytvořen z předpovědí na stejný čas, ale získaný v časových okamžicích $t, t+1$ a $t+2, \dots, t+k-1$ (viz obrázek 10). Výhodou časově zpožděných ensemblů je to, že nemusíme provádět dodatečné predikční výpočty (což v případě některých komplikovaných modelů může představovat značnou časovou úsporu), neboť obvykle předpovídáme na dostatečně dlouhý časový horizont a v každém kroku předpověď aktualizujeme. Podrobnější informace o ensemblovém předpovídání lze nalézt např. v [9].

8 Úsudkové (expertní) předpovídání

Lidský mozek i přes pokroky v oblasti neurověd a umělé inteligence je stále nejdokonalejším nástrojem, který můžeme používat ať již samostatně, anebo v kombinaci s podpůrnými nástroji pro jeho rozhodování. Jak již bylo naznačeno v úvodní části, předpověď chování komplexních systémů není jednoduchá úloha. I při využití sofistikovaných nelineárních predikčních metod vzniká potřeba předpovědi vhodným způsobem korigovat a korekce lidským expertem (experty) je jednou z neúčinnějších variant. Pochopitelně jsou situace, kdy úsudková predikce není možná (při řízení složitých zařízení v reálném čase, při potřebě předpovědi chování mnoha stovek, resp. tisíců veličin v minutových intervalech apod.).



Obrázek 10: Získání ensemblových předpovědí pomocí časově zpožděných předpovědí.

Predikce chování komplexních systémů pomocí lidských expertů se vyznačuje tím, že v naprosté většině nedochází k jednoznačným závěrům (viz například problematika globálního oteplování). Je to dáno složitostí vazeb a mnohdy i nejasností vztahů příčina – následek (viz studium chování makroekonomických veličin jako jsou HDP, nezaměstnanost, inflace apod.). Rozdílnost předpovědí je dána rozdílnou zkušeností jednotlivých expertů (každý využívá jinou učící množinu). Navíc stávající vazby, pokud existovaly v minulosti, se v průběhu času mění (svět se mění), a to mnohdy i dramaticky.

Zajímavá je také tendence myšlení k lineární extrapolaci, zatímco nelineární vazby mezi více proměnnými jsou velmi těžko pochopitelné (lidé mají rádi jednoduchá vysvětlení typu jednoduché implikace). Jeden z nejzásadnějších vlivů na úsudkové předpovědi chování komplexních systémů je problém očekávání (expektace) daného vývoje, které se projevuje podvědomým filtrováním získávaných dat tak, aby data odpovídala těmto očekáváním. Dokonce i výsledky velice sofistikovaných předpovědních modelů mohou být potlačovány či různým způsobem interpretovány tak, aby predikce nevybočily z očekávaného rámce. Toto je například jeden z možných důvodů špatné předpovědi ekonomické krize v posledních čtyřech letech, viz [6].

Názorná ilustrace vlivu očekávání na interpretaci obrazové informace je uvedena na obr. 11.



Obrázek 11: Dvojitá interpretace stejné vizuální informace – mladá dívka versus stará žena (zdroj internet).

Dalším jevem, který ovlivňuje expertní predikce, je příliš velké sebevědomí některých expertů anebo příliš malé sebevědomí a jejich snadná ovlivnitelnost hlasy ostatních expertů.

Přesto bývají expertní predikce velmi užitečné zvláště v kombinaci s matematickými modely, a to proto, že expert je intuitivně schopen použít pro predikce informace, jejichž užitečnost nemusí být na první pohled zřejmá a které v matematickém modelu nejsou zahrnuty (anebo je do modelu nebylo možné z ekonomických anebo technických důvodů zahrnout). Kromě toho je v určitých oblastech expertní predikce nezastupitelná (v lékařství) anebo je uživateli (např. veřejnou správou) vyžadována (např. v odborných stanoviscích).

9 Experiment s předpovědí globálního slunečního záření

Problém předpovědi počasí je problém úzce svázaný s problematikou modelování složitých přírodních dějů. Chaotická povaha systému popisujícího atmosférické jevy byla popsána v klasické práci E. Lorenze [10], která vedla ke zrodu nové vědecké disciplíny a k intenzivnímu studiu vlastností deterministického chaosu. Vzhledem k citlivosti vývoje počasí na počáteční podmínky je předpověditelnost počasí omezená. V současnosti se má za to, že teoretická mez predikčního horizontu pro předpověď počasí je okolo 3–4 týdnů [9]. Díky zvyšování výkonu počítačů, zpřesňování počátečních podmínek, zvyšování počtu a kvality pozorování a využívání moderních statistických metod, metod pro asi-

milaci dat a technik ensemblových předpovědí je dnes praktická doba předpověditelnosti počasí (tedy kdy je operativní předpověď lepší než klimatologická) okolo 15 dnů. Toto však neplatí pro meteorologické veličiny s regionálním charakterem (jako je i globální sluneční záření), kdy je praktický horizont předpověditelnosti mnohem nižší.

Předpověď globálního slunečního záření, ať již na krátkodobý horizont (1–72 hodin) anebo na dlouhodobý horizont (s využitím klimatologických hodnot) v jemném regionálním rozlišení je důležitou úlohou v souvislosti s prudkým nárůstem fotovoltaických instalací v Evropě i v ČR. Zatímco dlouhodobé předpovědi se používají k ekonomickým odhadům návratnosti investic do fotovoltaických farem, o krátkodobé předpovědi je zájem u společností zabývajících se výrobou a přenosem elektrické energie, u obchodníků s energiemi, u drobných provozovatelů fotovoltaických instalací i jinde.

Úloha předpovědi globálního slunečního záření není jednoduchá. Kromě lokálního charakteru může být problémem i to, že je potřeba predikovat hodnoty této veličiny s poměrně jemným časovým krokem (obvykle 1 hodina, ale i méně) a v jemném regionálním rozlišení (cca pod 10 km). Takováto předpověď není možná bez využití numerických předpovědních modelů počasí, které jsou označovány jako NWP (numerical weather prediction). Ty jsou založeny na numerickém řešení parciálních diferenciálních rovnic popisujících fyzikální procesy probíhající v atmosféře (dynamika proudění, bilance vodní páry, energetické toky apod.). Protože chování atmosféry je velmi komplikované, neexistuje universální nejlepší model počasí. Existují proto různé varianty NWP modelů, které se liší prostorovým i časovým měřítkem, detailností popisu fyzikálních jevů, idealizovanými modely proudění v terénu, různými přístupy pro modelování mikrofyziky oblaků apod. Při popisu dějů v atmosféře v NWP modelech se popisuje časový vývoj stavového vektoru obsahujícího prostorově diskretizované hodnoty fyzikálních veličin, jako jsou tlak, teplota, rychlost a směr větru, vlhkost, koncentrace dalších látek v ovzduší apod. Model popisuje vývoj stavů vždy vztažený k určité oblasti (doméně), která je rozdělena na řádově desítky vertikálních vrstev a horizontálně je pokrytá nejčastěji čtvercovou sítí. Vertikální členění je v NWP modelech nejčastěji řešeno podle měnicího se hydrostatického tlaku (v tzv. σ -hladinách). Časový krok pro modelování fyzikálních jevů se pohybuje v závislosti na horizontálním rozlišení okolo 10–90 sekund, uživatelské výstupy se obvykle poskytují s časovým krokem 1 hodina.

K předpovědi globálního slunečního záření jsme využili numerické předpovědní modely implementované v systému MEDARD [5], [12],

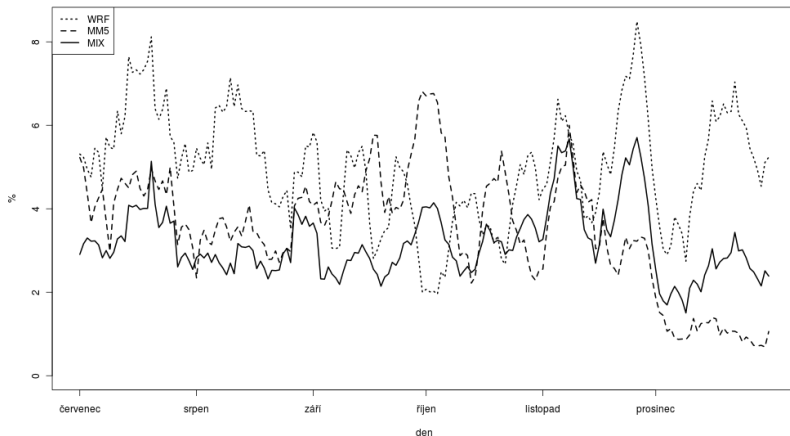
kteřý je operačně provozován v Ústavu informatiky AV ČR od r. 2004. Jádrem systému je NWP model WRF (Weather Research and Forecasting), který nahradil starší verzi modelu MM5 (Mesoscale Model, verze 5). Na jejich vývoji se podílela celá řada organizací, jako jsou National Center for Atmospheric Research (NCAR), National Oceanic and Atmospheric Administration (NOAA), National Centers for Environmental Prediction (NCEP) a další. Modely WRF i MM5 jsou zaměřeny především na vývoj regionálního počasí v rozlišení 2–20 km, experimentálně se zkoumají jeho vlastnosti v rozlišení 1 km a menším [8]. Okrajové podmínky jsou získávány z globálního modelu atmosféry GFS, který má v současné době vertikální rozlišení okolo 50 km. Ty jsou vystavovány v 6 hodinových intervalech, kdy je využito aktuální měření predikovaných stavů z předcházejícího kroku s využitím techniky asimilace dat. Zde se využívají informace z pozemních meteorologických stanic, radarových měření, balónových sond, měření na letadlech a lodích. Pro simulace v jemném rozlišení se počítá vnoření několika domén. V systému MEDARD se využívá tří vnořených domén, a to v rozlišení 27 km, 9 km a 3 km.

Výstupy z modelů MM5 a WRF byly použity jako vstupy do predikčního modelu pro předpověď výroby elektrické energie z fotovoltaických farem v ČR provozovaného v ČEPS a.s. Předpověď byla uskutečněna na předpovědní horizont „D+1“, tedy predikuje se výroba elektrické energie na následující den z rána (cca mezi 7–8 hodinou) dne aktuálního.

Vyhlazený průběh predikční chyby normované na instalovaný výkon je ukázán na obrázku 12. Z průběhu vidíme, že paradoxně starší verze NWP modelu MM5 vykazuje lepší výsledky než novější verze NWP modelu WRF. Je to dáno tím, že NWP modely nejsou apriorně navrhovány pro konkrétní (v našem případě energetické) potřeby, ale mají za úkol předpovídat počasí jako celek. Také se potvrzuje, že multimodelová předpověď (v našem případě je finální predikce počítána jako aritmetický průměr výstupů obou modelů a je na obrázku 12 označena jako MIX) může předpověď vylepšit.

Na obrázcích 13 a 14 je pak ukázka časově zpožděné ensemblové předpovědi (time-lagged ensemble) vytvořené postupně z 14 předpovědí NWP modelu MM5 v šestihodinovém kroku pro dva vybrané dny. Předpovědi jsou porovnány s měřením globálního slunečního záření na meteorologické stanici síť AIM (automatický imisní monitoring) provozované ČHMÚ. Zvýšená variabilita předpovědi se projevuje při zvýšené oblačnosti v odpoledních hodinách na obrázku 14. Časově zpožděné ensemblové předpovědi tak mohou být náhradou za jinak výpočetně

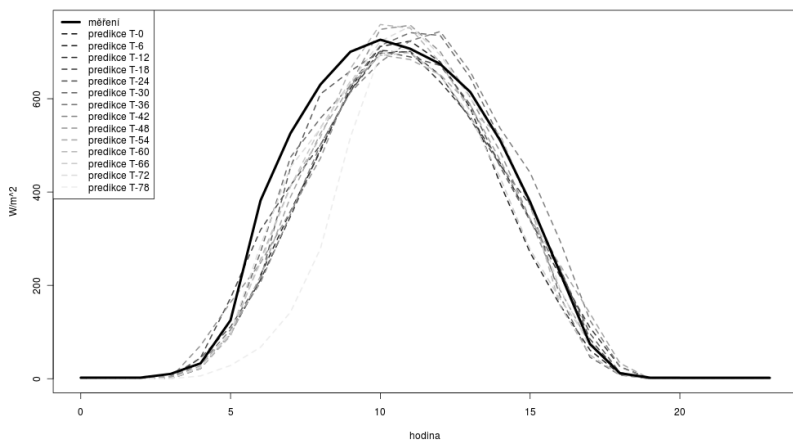
náročné klasické ensemblové předpovědi.



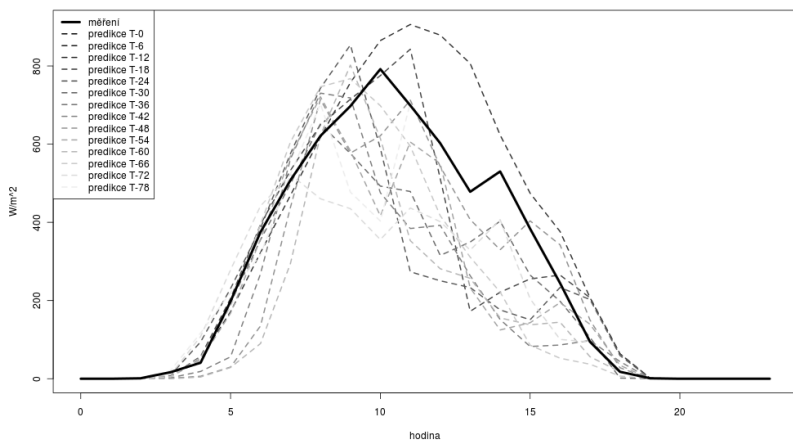
Obrázek 12: Průběh predikční chyby výroby elektrické energie z fotovoltaických zdrojů v ČR v druhé polovině roku 2011 (chyba je vyjádřena v procentech vzhledem k instalovanému výkonu) na predikční horizont D+1 (ranní předpověď na druhý den) s využitím numerických modelů počasí WRF a MM5 a jejich kombinace (MIX).

10 Závěrečná poznámka

Zájem o využití předpovědních metod v řadě oblastí lidské činnosti je stále velmi velký. Techniky předpovědi se stále zdokonalují v souladu s teoretickými i praktickými poznatky a s rozvojem výpočetní techniky, která umožňuje provádění dříve obtížně proveditelných experimentů a prodlužování predikčního horizontu s praktickou využitelností. Předpovědi poskytované sofistikovanými matematickými modely se staly součástí každodenního života. Velkou výzvou pro návrháře predikčních modelů je pak přesnější kvantifikace neurčitostí v předpovědích a zejména její srozumitelnější interpretace pro běžného uživatele.



Obrázek 13: Časově zpožděná ensamblová předpověď globálního slunečního záření na stanici Ondřejov pro den 15.5.2009.



Obrázek 14: Časově zpožděná ensamblová předpověď globálního slunečního záření na stanici Ondřejov se zvýšenou oblačností v odpoledních hodinách numerickým modelem počasí MM5 pro den 10.6.2008.

Literatura

- [1] Bishop Ch.: *Neural networks for Pattern Recognition*, Oxford University Press, 1996.
- [2] Boccarda N.: *Modeling Complex Systems*, Springer, New York, 2010.
- [3] Durbin J., Koopman S.J.: *Time Series Analysis by State Space Methods*, Oxford University press, Great Britain, 2004.
- [4] Eben K., Juruš P., Resler J., Belda M., Pelikán E., Krueger B.C., Keder J.: *An ensemble Kalman filter for short-term forecasting of tropospheric ozone concentrations*, Q.J.R. Meteorol. Soc., Vol. 131 (2005), pp. 3313-3322.
- [5] Eben, K., Juruš P., Resler J., Belda M., Krüger B.C., Keder J.: *MEDARD – an Environmental Modelling Project for the Territory of the Czech Republic*. ERCIM News, 61, 2005.
- [6] Fischer J.: *Macroeconomic Forecasts: Betraying Us When Most Needed*, keynote speech, International Symposium on Forecasting, Prague, June 2011
- [7] Grabec I., Sachse W.: *Synergetics of Measurement, Prediction and Control*, Springer Verlag, Berlin Heidelberg, 1997.
- [8] Juruš P.: *Modelování atmosféry v regionálním a lokálním měřítku s aplikacemi v dopravě*, disertační práce, Fakulta dopravní ČVUT, 2009.
- [9] Kalnay E.: *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, UK, 2004.
- [10] Lorenz, E. N.: *Deterministic nonperiodic flow*. Journal of the Atmospheric Sciences, 20, pp. 130–141. 1963.
- [11] Makridakis S., Wheelwright S.C., Hyndman R.J.: *Forecasting, Methods and Applications*, John Wiley & Sons, USA, 1998.
- [12] www.medard-online.cz
- [13] Nicolis G., Nicolis C.: *Foundations of Complex Systems*, World Scientific, Singapore, 2007.
- [14] Pelánek R.: *Modelování a simulace komplexních systémů*, Masarykova universita, Brno, 2011.

- [15] Pelikán E., de Groot C., Würtz D.: *Power Consumption in West-Bohemia: Improved Forecasts with Decorrelating Connectionist Networks*, Neural Network World, No.6, Vol.2, 1992, pp.701-712.
- [16] Pena D., Tiao G.C., Tsay R.S.: *A course in Time series Analysis*, John Wiley & Sons, Canada, 2001.
- [17] Raidl A.: *O (ne)předpověditelnosti počasí*, Pokroky matematiky, fyziky, a astronomie, Vol. 42, No.6, pp.302-312, 1997.
- [18] Resler J.: *Modelování atmosféry v regionálním a lokálním měřítku s aplikacemi v dopravě*, disertační práce, Fakulta dopravní ČVUT, 2009.
- [19] Scholkopf B., Smola A.J.: *Learning with Kernels, Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, Massachusetts, 2002
- [20] Takens F.: *Detecting strange attractors in turbulence*. In D. A. Rand and L.-S. Young. *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics, vol. 898. Springer-Verlag. pp. 366–381, 1981

11 Doc. Ing. Emil Pelikán, CSc.

Vzdělání

1974–1979 vysokoškolské, obor matematické inženýrství, FJFI ČVUT

1985 obhájena kandidátská disertační práce „Počítačové zpracování kvaziperiodických signálů“, Ústav informatiky a výpočetní techniky ČSAV

2002 obhájena docentská habilitační práce, obor inženýrská informatika, ČVUT Fakulta dopravní

Zahraniční stáže

1992 Swiss Centre for Supercomputing, ETH Zuerich, Švýcarsko

1998 Issac Newton Institute, Cambridge , Velká Británie

Zaměstnání

1979 – dosud Ústav informatiky AV ČR (dříve Ústav informatiky a výpočetní techniky ČSAV, dříve Středisko výpočetní techniky ČSAV), 1979–1985 aspirantské studium, 1985–1993 vědecký pracovník, od r. 1993 vedoucí vědecký pracovník (1. pracovní poměr)

1998 – dosud ČVUT Fakulta dopravní, vysokoškolský učitel (2. pracovní poměr)

Pedagogické a výzkumné zaměření

Výzkum v oblasti modelování složitých procesů s aplikacemi v environmentální informatice, energetice, bankovníctví

1998 – dosud výuka předmětu „Predikční metody“, ČVUT Fakulta dopravní, Praha

Publikační činnost a další důležité odborné aktivity

Doc. Ing. Emil Pelikán, CSc. je autorem více než 120 odborných a vědeckých prací z oblasti analýzy časových řad, umělých neuronových sítí, modelování složitých procesů a predikčních metod, které vyšly v našich i zahraničních časopisech anebo byly publikovány ve sbornících významných mezinárodních konferencí. Přednesl řadu zvaných přednášek, mimo jiné na universitách v USA, Velké Británii, Švýcarsku, Portugalsku a v Brazílii. Vede vědecké oddělení nelineárního modelování a přímo se podílí na vedení odborných týmů zabývajících se interdisciplinární problematikou modelování složitých procesů s dopady v řadě aplikačních oblastí. Doc. Emil Pelikán byl hlavním navrhovatelem anebo spolunavrhovatelem více než desítky grantových projektů podporovaných EU (5RP), GA ČR, GA AV ČR, MŽP a aplikovanými projekty podporovanými např. ČHMÚ, Moravskoslezským krajem, ČEZ, RWE apod.

Byl členem mezinárodního panelu NATO (Science for Peace program, Information and Security panel), je členem odborné společnosti International Institute of Forecasters, byl též členem redakční rady amerického časopisu International Journal of Forecasting. Účastní se výuky na FD ČVUT, je členem oborových rad ČVUT a člen programových i organizačních výborů mezinárodních konferencí.

Vybrané publikace

- Brabec, M.; Malý, M.; Pelikán, E.; Konár, O. *Statistical Model of Segment-Specific Relationship Between Natural Gas Consumption and Temperature in Daily and Hourly Resolution*. In Natural Gas. Rijeka: Sciyo, 2010. s. 393–416. ISBN 978-953-307-112-1.
- Pelikán, E.; Eben, K.; Resler, J.; Juruš, P.; Krč, P.; Brabec, M.; Brabec, T.; Musílek, P. *Wind power forecasting by an empirical model using NWP outputs*. In Environment and Electrical Engineering. IEEE, 2010. s. 45–48. ISBN 978-1-4244-5370-2.
- Vejmelka, M.; Musílek, P. ; Paluš, M.; Pelikán, E. *K-Means Clustering for Problems with Periodic Attributes*. International Journal of Pattern Recognition and Artificial Intelligence, 2009, Roč. 23, č. 4, s. 721–743. ISSN 0218-0014
- Brabec, M.; Konár, O.; Malý, M.; Pelikán, E.; Vondráček, J. *A Statistical Model for Natural Gas Standardized Load Profiles*. Journal of the Royal Statistical Society Series C-Applied Statistics, 2009, Roč. 58, č. 1, s. 123–139. ISSN 0035-9254
- Brabec, M.; Konár, O.; Pelikán, E.; Malý, M. *A Nonlinear Mixed Effects Model for the Prediction of Natural Gas Consumption by Individual Customers*. International Journal of Forecasting, 2008, Roč. 24, č. 4, s. 659–678. ISSN 0169-2070.
- Vondráček, J.; Pelikán, E.; Konár, O.; Čermáková, J.; Eben, K.; Malý, M.; Brabec, M. *A Statistical Model for the Estimation of Natural Gas Consumption*. Applied Energy, 2008, Roč. 85, č. 5, s. 362–370. ISSN 0306-2619.
- Šimůnek, M.; Pelikán, E. *Temperatures Data Preprocessing for Short-Term Gas Consumption Forecast*. In IEEE International Symposium on Industrial Electronics ISIE'08. IEEE, 2008. s. 1192–1196. ISBN 978-1-4244-1666-0.
- Schlink, U.; Herbath, O.; Richter, M.; Dorling, S.; Nunnari, G.; Cawley, G.; Pelikán, E. *Statistical Models to Assess the Health Effects and to Forecast Ground Level Ozone*. Environmental Modelling & Software, 2006, Roč. 21, č. 4, s. 547–558. ISSN 1364-8152.

Pelikán, E.; Šimůnek, M. *Risk Management of the Natural Gas Consumption using Genetic Algorithms*. Neural Network World, 2005, Roč. 15, č. 5, s. 425–436. ISSN 1210-0552.

Eben, K.; Juruš, P.; Resler, J.; Belda, M.; Pelikán, E.; Krüger, B.C.; Keder, J. *An Ensemble Kalman Filter for Short-Term Forecasting of Tropospheric Ozone Concentrations*. Quarterly Journal of the Royal Meteorological Society, 2005, Roč. 131, č. 613, s. 3313–3322. ISSN 0035-9009.

Svítek, M.; Pelikán, E. *Monitoring and Control of Dangerous Goods Transport*. Neural Network World, 2004, Roč. 14, č. 3-4, s. 303-312. ISSN 1210-0552.

Schlink, U.; Dorling, S.R.; Pelikán, E.; Nunnari, G.; Cawley, C.; Junninen, H.; Greig, A.; Foxall, R.; Eben, K.; Chatterton, T.; Vondráček, J.; Richter, M.; Dostál, M.; Bertuccio, L.; Kolehmainen, M.; Doyle, M. *A Rigorous Inter-Comparison of Ground-Level Ozone Predictions*. Atmospheric Environment, 2003, Roč. 37, č. 23, s. 3237–3253. ISSN 1352-2310.

Pelikán, E.; Eben, K.; Vondráček, J.; Dostál, M.; Krejčíř, P.; Keder, J. *On the Typology of Daily Courses of Tropospheric Ozone Concentrations*. Wien: Springer, 2001. ISBN 3-211-83651-9. In Artificial Neural Nets and Genetic Algorithms. Proceedings of the International conference s. 461–464.

Pelikán, E.; Eben, K.; Vondráček, J.; Krejčíř, P.; Keder, J. *Ground Level Ozone Peak Forecast using Neural Networks and Kalman Filter*. Meteorological Journal, 2000, Roč. 3, č. 2, s. 3–8. ISSN 1335-339X.

Pelikán, E.; Petrák, L. *Electric Load Modelling with Theory of Extreme Values*. Neural Network World, 1999, Roč. 9, č. 4, s. 297–304. ISSN 1210-0552.

Pelikán, E. *Principles of Forecasting - A Short Overview*. In SOF-SEM'99: Theory and Practice of Informatics. Berlin : Springer, 1999. s. 311–327. ISBN 3-540-66694-X. ISSN 0302-9743.

Pelikán, E.; Eben, K.; Petrák, L. *One-Day Prediction of Electric Load Reflecting Future RCS Schedule*. Journal of Forecasting, 1996, Roč. 15, č. 6, s. 427–435. ISSN 0277-6693.

Khanna, V.; Štáva, E.; Pelikán, E. *Estimating of CZK Volatility*. Neural Network World, 1995, Roč. 5, č. 6, s. 905–910. ISSN 1210-0552.

Pecen, L.; Ramešová, N.; Pelikán, E.; Beran, H. *Application of the GUHA Method on Financial Data*. Neural Network World, 1995, Roč. 5, č. 4, s. 565–571. ISSN 1210-0552.

Pelikán, E.; Groot, de C.; Wuertz, D. *Power Consumption in West-Bohemia: Improved Forecasts with Decorrelating Connectionist Networks*. Neural Network World, 1992, Roč. 2, č. 6, s. 701–711. ISSN 1210-0552.

Pelikán, E. *Spectral analysis of ARMA processes by Prony's method*. Kybernetika, 1984, Roč. 20, č. 4, s. 322. ISSN 0023-5954.

Pelikán, E.; Vosvrda, M. *On covariance coefficients estimates of finite order moving average processes*. Kybernetika, 1981, Roč. 17, č. 2, s. 169. ISSN 0023-5954.