

Univerza v Ljubljani
Fakulteta za elektrotehniko

Kristjan Ažman

**Identifikacija dinamičnih sistemov z Gaussovimi
procesimi z vključenimi linearnimi modeli**

Magistrsko delo

Mentor: prof. dr. Juš Kocijan

V Ljubljani, september 2004

Zahvala

Rad bi se zahvalil mentorju prof. dr. Jušu Kocijanu za spodbudo, ideje in potrpežljivost ter Andreju, Gregorju, Miru in Janezu za pomoč pri delu z modelno napravo. Hvala tudi prof. dr. Dragu Matku za večletno podporo.

Povzetek

Identifikacija dinamičnih sistemov z modeli na osnovi *Gaussovih procesov* (GP) je relativno novo področje raziskovanja. Ta pristop je verjetnostna nelinearna identifikacija na principu "črne škatle" in je soroden metodam identifikacije z umetno nevronske mreže ali mehki modeli. Modeli na osnovi Gaussovih procesov imajo zanimivo in za vodenje uporabno lastnost, da model poleg vrednosti izhoda napoveduje tudi zaupanje v to vrednost. To je pomembno, saj pove v kolikšni meri lahko zaupamo predikciji modela za določen vhodni signal.

Namen magistrskega dela je seznanitev z identifikacijo nelinearnih dinamičnih sistemov z Gaussovimi procesi, vpeljava predznanja v Gaussove procese v obliki lokalnih linearnih modelov ter ovrednotenje metode na izbranih primerih nelinearnih dinamičnih sistemov.

V magistrskem delu je najprej predstavljen osnovni princip identifikacije z Gaussovimi procesi, ki je ponazorjen s preprostimi primeri. Nato je predstavljena razširitev GP s porazdelitvijo kot vhodom v GP. Predikcija identificiranih GP modelov je vseskozi blizu pravi vrednosti ali pa nas na poslabšanje napovedi opozori povečana varianca. Relativno veliko število točk nelinearnega modela lahko zmanjšamo z vključevanjem linearnih lokalnih modelov. Identifikacijo takega modela smo preizkusili na dinamičnem sistemu prvega reda in na dinamičnem sistemu drugega reda. Razširjanje negotovosti skozi GP model z vključenimi lokalnimi modeli smo lahko izvedli samo za sisteme prvega reda.

Metodo identifikacije brez in z vključenimi lokalnimi modeli smo uporabili tudi na podatkih meritev modelne naprave.

Metoda identifikacije z Gaussovimi procesi ima kar nekaj potencialnih prednosti pred podobnimi metodami za identifikacijo (Takagi-Sugeno mehki modeli, umetne nevronske mreže). Glavna prednost in razlika je, da poleg napovedi daje

tudi zaupanje v to napoved, kar se lahko koristno uporabi pri vodenju sistema. Druga razlika je, da je potrebno optimizirati precej manjše število parametrov kot npr. pri nevronskih mrežah, kar pomeni lažjo optimizacijo z boljšimi možnostmi, da se konča v globalnem minimumu. Kot negativno lastnost metode lahko navedemo veliko računsko zahtevnost pri optimizaciji hiperparametrov, ki raste s tretjo potenco velikosti učne množice. Z vključevanjem linearnih lokalnih modelov v metodo identifikacije z Gaussovimi procesi lahko nadomestimo veliko število sicer potrebnih učnih točk in tako zmanjšamo računsko breme. Verjetno še pomembnejša prednost vključevanja lokalnih modelov je možnost vključitve lokalne informacije v globalno ne da bi se informacija lokalnega značaja izgubila.

Posledica dejstva, da je metoda še v nastajanju, je precej odprtih vprašanj. Odprto vprašanje je razširjanje negotovosti skozi GP model dinamičnega sistema višjega reda z vključevanjem lokalnih modelov. Potrebno bi se bilo posvetiti tudi učinkovitosti programske opreme za identifikacijo GP modelov.

Pristop kaže na vrsto možnih uporab na področju analize in načrtovanja sistemov vodenja in obeta zanimive raziskovalne probleme.

Abstract

Gaussian process model identification of dynamical systems is relatively new field of research. It is a member of probabilistic non-parametric black-box identification methods and is comparable to more frequently used fuzzy models or artificial neural networks. The GP model has unique and interesting quality, that besides predicting the output of the identified system, the model also provides the level of confidence in its prediction for certain input.

The purpose of this thesis is to describe the principle of GP modelling of nonlinear dynamic systems, to introduce the information into the GP model in the form of linear local models and to evaluate the GP identification model on examples of nonlinear dynamic systems.

First the basic principles of GP model identification is presented together with some simple examples. GP model's expansion, where complete distributions of past outputs are used as inputs, is also realized and compared to basic GP model. Relatively large number of points of non-parametric model can be effectively reduced by incorporating linear local models. Identification of this model was evaluated on dynamical systems of first and second order. Uncertainty propagation through GP model could only be realized for first order dynamical systems.

The method of identification, with and without incorporation of local models, was then applied on measurement data from laboratory pilot plant.

GP model identification has certain potential benefits when compared to other similar identification methods (Takagi-Sugeno fuzzy models, artificial neural networks). Its main difference and advantage is that it provides the most probable output of the system together with the level of confidence it has about the output, which can be useful for controlling the system. The other advantage is a smaller number of used parameters, which reduces the problem of optimization

stuck in local minimum. Its disadvantage is computational burden associated with inverting covariance matrix, which has to be calculated in every step of the optimization. This can be effectively reduced by introducing prior information in form of linear local models into GP model. Beside reducing the number of data used for learning, local models modelling data near equilibrium curve can effectively model the main attributes of the systems and are relatively easy to identify. It also provides a chance to include local information about the system into global model without loss of information.

The fact that method is still in nascent stage leaves us with quite a few unanswered questions. An example of it is propagation of uncertainty throughout the GP model with incorporated local models of dynamic systems of higher order. Performance of GP identification software should also be improved.

The GP model approach shows many potential applications for system analysis and control design and promises interesting research problems.

Vsebina

1. Uvod	1
2. Gaussovi procesi	3
2.1 Osnovni pojmi	3
2.2 Modeliranje z Gaussovimi procesi	4
2.3 Kovariančna funkcija	4
2.4 Optimizacija	5
2.5 Napovedovanje	6
2.6 Primer uporabe GP za identifikacijo statičnega nelinearnega sistema	8
3. Uporaba Gaussovih procesov za identifikacijo nelinearnih dinamičnih sistemov	13
3.1 Identifikacija dinamičnih sistemov	13
3.2 Primera identifikacije nelinearnih dinamičnih sistemov	14
3.3 Predikcija pri napovedani srednji vrednosti kot vhodu v GP model	16
3.3.1 Dinamični sistem prvega reda	16
3.3.2 Dinamični sistem drugega reda	20
3.4 Predikcija pri porazdelitvi kot vhodu v GP model	26
3.5 Primerjava med modeli	30
3.5.1 Dinamični sistem prvega reda	30
3.5.2 Dinamični sistem drugega reda	32

4. Vključevanje linearnih lokalnih modelov v GP modele	37
4.1 Kovariančna funkcija	39
4.2 Predikcija pri porazdelitvi kot vhodu v LM GP model	40
4.3 Napovedana srednja vrednost za nov $\mathbf{x} \sim \mathcal{N}_{\mathbf{x}}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$	41
4.4 Napovedana varianca za nov $\mathbf{x} \sim \mathcal{N}_{\mathbf{x}}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$	42
4.5 Simulacija	43
4.6 Programska oprema za identifikacijo z vključenimi lokalnimi modeli	43
4.7 Primera vključevanja lokalnih modelov v GP model	44
4.7.1 Dinamični sistem prvega reda	44
4.7.2 Dinamični sistem drugega reda	54
5. Primer identifikacije modelne naprave z GP	59
5.1 Opis naprave	59
5.1.1 Seznam senzorjev	62
5.1.2 Seznam aktuatorjev	62
5.1.3 Uporabljeni podsistem naprave	62
5.1.4 Matematični model naprave	64
5.2 Opis področja delovanja naprave in meritev	65
5.3 Identifikacija GP modela	67
5.4 Identifikacija GP modela z vključenimi lokalnimi modeli	73
5.5 Primerjava rezultatov in komentar	76
6. Zaključek	77
Literatura	81

Seznam slik

2.1	Nelinearna funkcija	8
2.2	Primerjava identifikacije z metodo najmanjših kvadratov (LSQ) in z GPi	10
2.3	Primerjava identifikacije z metodo LSQ z znano strukturo funkcije in identifikacije z GPi	10
2.4	Primerjava napake predikcije metod LSQ in GP	11
3.1	Nelinearni dinamični sistem prvega reda in učne točke	14
3.2	Statična karakteristika nelinearnega dinamičnega sistema drugega reda	15
3.3	Normiran odziv na stopnico v različnih točkah nelinearnega dinamičnega sistema drugega reda	16
3.4	Signala za učenje in vrednotenje – dinamični sistem prvega reda	17
3.5	Napovedi GP modela in nelinearnost originalnega sistema	18
3.6	Napaka GP	18
3.7	Varianca GP	19
3.8	Simulacija na testnem signalu	19
3.9	Avtokorelacija napake $\Phi_{ee}(\tau)$	20
3.10	Križna korelacija med vzbujevalnim vhodnim signalom in napako $\Phi_{ue}(\tau)$	21
3.11	Signal za učenje dinamičnega sistema drugega reda	22
3.12	Signal za vrednotenje dinamičnega sistema drugega reda	22
3.13	Simulacija GP modela za nelinearni dinamični sistem drugega reda	23

3.14	Absolutna napaka $ e(k) $ simulacije GP modela, nelinearni dinamični sistem drugega reda	24
3.15	Avtokorelacija napake predikcije GP modela Φ_{ee} , nelinearni dinamični sistem drugega reda	24
3.16	Križna korelacija vhoda in napake predikcije GP modela Φ_{ue} , nelinearni dinamični sistem drugega reda	25
3.17	Simulacija nelinearnega dinamičnega sistema z GP modelom	29
3.18	Primerjava GPov z in brez uporabe variance na vhodu	31
3.19	Primerjava GPov z in brez uporabe variance na vhodu - razširjeno področje	31
3.20	Avtokorelacija napake Φ_{ee}	32
3.21	Križna korelacija med vhomom in napako Φ_{ue}	33
3.22	Simulacija "eksaktnega" GP modela dinamičnega sistema drugega reda	33
3.23	Napaka predikcije eksaktnega GP modela skupaj z zaupanjem	34
3.24	Avtokorelacija napake Φ_{ee}	34
3.25	Križna korelacija med vhomom in napako Φ_{ue}	35
4.1	Učni podatki	44
4.2	Signal za vrednotenje LM GP modela za din. sistem prvega reda	45
4.3	Enokoračna predikcija (samo lokalni modeli)	47
4.4	Napaka enokoračne predikcije (samo lokalni modeli)	47
4.5	Varianca (samo lokalni modeli)	48
4.6	Enokoračna predikcija	48
4.7	Napaka enokoračne predikcije	49
4.8	Varianca	49
4.9	Simulacija (samo lokalni modeli)	50
4.10	Simulacija z "naivno" metodo	51
4.11	Simulacija z "naivno" in "eksaktno" metodo	51

4.12	Napaka simulacije GP pri "naivnem" in "eksaktnem" pristopu . . .	52
4.13	Primerjava variance "naivne" in "eksaktne" metode na odseku signala, $u(k)$ povečan zaradi ilustrativnosti	52
4.14	Avtokorelacija napake $\Phi_{ee}(\tau)$	53
4.15	Križna korelacija med vzbujevalnim vhodnim signalom in napako $\Phi_{ue}(\tau)$	53
4.16	Simulacija GP modela za nelinearni dinamični sistem drugega reda	55
4.17	Absolutna napaka $ e(k) $ simulacije GP modela za nelinearni dinamični sistem drugega reda	56
4.18	Avtokorelacija napake predikcije GP modela Φ_{ee} za nelinearni dinamični sistem drugega reda	56
4.19	Križna korelacija vhoda in napake predikcije GP modela Φ_{ue} za nelinearni dinamični sistem drugega reda	57
5.1	Laboratorijska modelna naprava treh tankov	60
5.2	Procesna shema modelne naprave treh tankov	61
5.3	Procesna shema uporabljenega podsistema	64
5.4	Statična karakteristika sistema	66
5.5	Signal za identifikacijo lokalnega modela v delovni točki $(U_{eq}, Y_{eq}) = (3, 23.5)$	67
5.6	Signal za vrednotenje lokalnega modela v delovni točki $(U_{eq}, Y_{eq}) = (3, 23.5)$	68
5.7	Signal za učenje, najhitrejša sprememba v času $T = 10s$	68
5.8	Signal za vrednotenje, najhitrejša sprememba v času $T = 20s$	69
5.9	Simulacija GP modela	70
5.10	Napaka simulacije GP modela	70
5.11	Avtokorelacija napake GP modela	71
5.12	Križna korelacija napake GP modela in vhodnega signala u	71
5.13	Simulacija "eksaktnega" GP modela	72

5.14	Napaka simulacije "eksaktnega" GP modela	72
5.15	Simulacija GP modela	74
5.16	Napaka simulacije GP modela	74
5.17	Avtokorelacija pogreška GP modela z vključenimi LM	75
5.18	Križna korelacija pogreška simulacije GP modela z vključenimi LM	75

Seznam tabel

4.1	Parametri uporabljenih lokalnih modelov za dinamični sistem prvega reda	46
4.2	Parametri uporabljenih lokalnih modelov za nelinearni dinamični sistem drugega reda	54
5.1	Tabela spremenljivk modela	63
5.2	Tabela aktuatorjev modela	63
5.3	Tabela konstantnih parametrov modela	65
5.4	Parametri uporabljenih lokalnih modelov za model treh tankov . .	73

1. Uvod

Identifikacija dinamičnih sistemov z modeli na osnovi *Gaussovih procesov* (GP) je relativno novo področje raziskovanja. Ta pristop je verjetnostna neparametrična identifikacija na principu "črne škatle" in je soroden metodam identifikacije z umetno nevronske mrežo ali mehki modeli. Modeli na osnovi Gaussovih procesov imajo zanimivo in za vodenje uporabno lastnost, da model poleg vrednosti izhoda napoveduje tudi zaupanje v to vrednost. To je pomembno, saj pove v kolikšni meri lahko zaupamo predikciji modela za določen vhodni signal.

Namen magistrskega dela je seznanitev z identifikacijo nelinearnih dinamičnih sistemov z Gaussovimi procesi, vpeljava predznanja v Gaussove procese v obliki lokalnih linearnih modelov ter ovrednotenje metode na izbranih primerih nelinearnih dinamičnih sistemov.

Gaussovi procesi so na področju statistike znani že zelo dolgo. Primer njihove uporabe lahko najdemo na področjih kot je npr. geostatistika, kjer so GP modeli znani pod imenom "kriging" [5]. Podrobnejše informacije o zgodovini uporabe lahko najdemo v [13]. Prvi je GP za reševanje regresijskega problema uporabil O'Hagan[19] leta 1978, vendar njegovi rezultati niso vzbudili večje pozornosti. GPi so pridobili popularnost predvsem na področju strojnega učenja z deli Neala (npr. [18]) in Rasmussena (npr. [21]) od leta 1994 naprej. Leta 1996 je bil GP model predstavljen v okviru Bayesove teorije.

Povezavo GP modelov z nekaterimi drugimi pristopi kot so Gaussove nevronske mreže, zlepi in nekatere druge metode najdemo v [13].

Uporaba GP modelov za identifikacijo nelinearnih dinamičnih sistemov pa je bila eno izmed glavnih področij evropskega projekta MAC, katerega partner je bil tudi Odsek za sisteme in vodenje Inštituta Jožef Stefan in je potekal od leta 2000 do 2004. S tem projektom so se pričela pojavljati tudi prva dela, ki

opisujejo GP modele dinamičnih sistemov (npr. [5, 3, 6, 9, 10]). Ideja, kako lahko v neparametrične modele vključimo dinamične linearne lokalne modele je bila prvič predstavljena leta 2002 v delih [11] in [22].

Do sedaj še nismo zasledili objavljenih rezultatov vključevanja linearnih lokalnih modelov v GP modele za dinamične sisteme več kot prvega reda oz. vrednotenje te metode na meritvah, kar je namen te naloge.

Magistrsko delo ima naslednjo zgradbo. V naslednjem poglavju sledi kratek opis Gaussovih procesov in osnovnih pojmov ter primer uporabe metode na statičnem problemu. V tretjem poglavju je GP model uporabljen za identifikacijo dveh nelinearnih dinamičnih sistemov. Tu je opisana tudi razširitev GP modela, kjer kot vhod vanj nastopa (predpostavljena) normalna porazdelitev iz izhoda modela. Naslednje poglavje opisuje razširitev GP modela z lokalnimi modeli ter prikaže uporabo na nelinearnih dinamičnih sistem prvega in drugega reda. Peto poglavje opisuje identifikacijo modelne naprave z Gaussovimi procesi. Zaključek povzema rezultate, izpostavi glavne ugotovitve ter daje smernice za prihodnje raziskovanje.

2. Gaussovi procesi

V tem poglavju bodo predstavljeni osnovni pojmi, povezani z Gaussovimi procesi. Več o njih najdemo v [2] in [12].

2.1 Osnovni pojmi

Naključna spremenljivka je spremenljivka, katere vrednost (realizacija) je odvisna od naključja. Naključni proces je večkratna realizacija naključne spremenljivke v odvisnosti od nekih (neke) neodvisnih spremenljivk (npr. časa, prostora). Če so realizacije te spremenljivke porazdeljene po normalnem (Gaussovem) porazdelitvenem zakonu, takemu procesu rečemo *Gaussov proces* [2, 12]. Ta je povsem določen s srednjo vrednostjo $\boldsymbol{\mu}(\mathbf{X})$ in kovariačno matriko $\boldsymbol{\Sigma}_{\mathbf{x}}(\mathbf{X})$, kjer je \mathbf{X} matrika vhodnih vektorjev neodvisnih spremenljivk \mathbf{x}_i za posamezno realizacijo y_i in i zaporedno število vhodnega vektorja.

Predstavljajmo si regresijski problem. Imamo nabor vhodnih vektorjev, združenih v matriko \mathbf{X} in nabor izhodnih točk \mathbf{y} . Želimo najti funkcijo $f(\mathbf{x})$, ki naj čimbolje opisuje relacijo med N vhodno-izhodnimi pari (\mathbf{x}_i, y_i) .

Modeliranje z Gaussovimi procesi se ravna po Bayes-ovem principu modeliranja, kjer namesto parametrizacije funkcije znanje uporabimo za določitev porazdelitvenega zakona nad neko družino funkcij, ki preslikavajo vhodne podatke \mathbf{x}_i v izhodne y_i . Predhodno znanje (ang. prior) je vključeno tako, da odraža mnenje o preslikavi med vhodi na izhodi in običajno predpostavlja gladkost preslikave ("podobni" vhodi naj bi se preslikali v "podobne" izhode). Ko vključimo v model še verjetnost učne množice (ang. likelihood), sestavljene iz N vhodno-izhodnih parov (\mathbf{x}_i, y_i) , dobimo "a-posteriori" porazdelitev za predikcijo modela.

2.2 Modeliranje z Gaussovimi procesi

Enostaven porazdelitveni zakon, ki velja za družino funkcij in ga lahko uporabimo, je normalna (Gaussova) porazdelitev. Na Gaussov proces lahko gledamo kot na množico naključnih spremenljivk, ki so porazdeljene po normalni porazdelitvi¹ [2]: $f(\mathbf{x}_1), \dots, f(\mathbf{x}_N) \sim \mathcal{N}(0, \mathbf{K})$, $\mathbf{K} = [K_{ij}]$, kjer je K_{ij} od vhodnih vektorjev \mathbf{x}_i in \mathbf{x}_j odvisna kovarianca med izhodoma $f(\mathbf{x}_i)$ in $f(\mathbf{x}_j)$. Gaussov proces je povsem določen s srednjo vrednostjo in kovariančno matriko \mathbf{K} , katere elementi se računajo s kovariančno funkcijo [2].

2.3 Kovariančna funkcija

Vrednost kovariančne funkcije $C(\mathbf{x}_i, \mathbf{x}_j)$ izraža korelacijo med posameznima izhodoma $f(\mathbf{x}_i)$ in $f(\mathbf{x}_j)$ glede na vhoda \mathbf{x}_i in \mathbf{x}_j . Kovariančna funkcija je lahko katerakoli funkcija, ki tvori ne-negativno definitno kovariančno matriko \mathbf{K} za katerikoli nabor vhodnih vektorjev \mathbf{X} . Ponavadi je izbrana taka kovariančna funkcija, da so točke, ki so bližje skupaj v vhodnem prostoru, tudi močnejše korelirane. Pri predpostavki o stacionarnosti procesa², je najpogosteje uporabljena Gaussova kovariančna funkcija, kjer element K_{ij} kovariančne matrike \mathbf{K} izračunamo kot [2]:

$$K_{ij} = C(\mathbf{x}_i, \mathbf{x}_j) = v \exp \left[-\frac{1}{2} \sum_{d=1}^D w_d (x_i^d - x_j^d)^2 \right] \quad (2.1)$$

oz. v vektorski obliki:

$$C(\mathbf{x}_i, \mathbf{x}_j) = v \exp \left[-\frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{W}^{-1} (\mathbf{x}_i - \mathbf{x}_j) \right] \quad (2.2)$$

in

$$\mathbf{W}^{-1} = \begin{bmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_n \end{bmatrix} \quad (2.3)$$

¹Prepostavimo sistem s srednjo vrednostjo nič, kar lahko vedno dosežemo s primernim normiranjem.

²Proces je stacionaren, če je kovarianca med dvema točkama odvisna samo od medsebojne razdalje in neodvisna od premika v prostoru.

Parametri v, w_1, w_2, \dots, w_D so svobodno določljivi parametri kovariančne funkcije - *hiperparametri* [21] in izražajo predhodno znanje (prior). D je dolžina vhodnega vektorja \mathbf{x} .

Da dana kovariančna matrika ustreza pogojem ne-negativne definitnosti, morajo biti vsi hiperparametri večji od nič. Parameter v skalira velikost variance, parametri w_i pa odražajo relativno pomembnost posamezne komponente vhodnega signala.

2.4 Optimizacija

Predpostavimo vhodno/izhodno relacijo $y = f(\mathbf{x}) + \epsilon$, kjer je na izhodu sistema dodan beli šum z normalno porazdelitvijo in varianco v_0 , kar označimo z $\epsilon \sim \mathcal{N}(0, v_0)$. Prepostavimo tudi normalno porazdelitev funkcije $f(\cdot)$ s kovariančno funkcijo (2.1) z neznanimi hiperparametri. Pri tej predpostavki za kovariančno matriko velja $y_1, \dots, y_n \sim \mathcal{N}(0, \mathbf{K}_N)$, kjer $K_{N_{ij}} = \Sigma_{ij} + v_0\delta_{ij}$, kjer je $\delta_{ij} = 1$, če $i = j$ in $\delta_{ij} = 0$ v vseh ostalih primerih.

Na podlagi N -tih parov \mathbf{x}_i, y_i želimo napovedati porazdelitev izhoda y^* pri vходу x^* . Lahko zapišemo [2] $\mathbf{y}, y^* \sim \mathcal{N}(0, \mathbf{K}_{N+1})$, kjer

$$\mathbf{K}_{N+1} = \begin{bmatrix} \begin{bmatrix} \mathbf{K} \\ \mathbf{k}(\mathbf{x}) \end{bmatrix} \\ \begin{bmatrix} \mathbf{k}(\mathbf{x})^T \\ k(\mathbf{x}) \end{bmatrix} \end{bmatrix} \quad (2.4)$$

Skupno verjetnost odziva lahko razdelimo na dva dela. Pogojni del (ang. conditional part) napoveduje naš odziv za \mathbf{x}^* , če poznamo del, ki določa verjetnost naših ciljnih vrednosti glede na vhode (ang. marginal part): $p(\mathbf{y}|\mathbf{X}) \sim \mathcal{N}(0, \mathbf{K})$. Tu je \mathbf{y} vektor učnih izhodov velikosti $N \times 1$ in \mathbf{X} matrika učnih vhodov velikosti $N \times D$.

Drugače predstavljeno: naloga zahteva izračun verjetnosti odziva [12]:

$$p(y^*|\mathbf{x}^*, \mathbf{y}, \mathbf{X}) = \int p(y^*|\mathbf{x}^*, \boldsymbol{\Theta}, \mathbf{y}, \mathbf{X})p(\boldsymbol{\Theta}|\mathbf{y}, \mathbf{X})d\boldsymbol{\Theta} \quad (2.5)$$

Ker je ta integral ponavadi analitično neizračunljiv, obstajata dve alternativni [12].

Prva alternativa je, da integral aproksimiramo z uporabo najbolj verjetnih vrednosti hiperparametrov. Naša naloga je oceniti vrednost hiperparametrov kovariančne funkcije in vrednost variance šuma v_0 . To naredimo tako, da poiščemo največjo vrednost logaritma porazdelitve danih vhodno-izhodno parov (ang. log-likelihood):

$$\mathcal{L}(\boldsymbol{\Theta}) = \log(p(\mathbf{y}|\mathbf{X})) = -\frac{1}{2} \log(|\mathbf{K}|) - \frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} - \frac{N}{2} \log(2\pi) \quad (2.6)$$

kjer je $\boldsymbol{\Theta}$ vektor parametrov, $\boldsymbol{\Theta} = [w_1 \dots w_D v v_0]^T$ in \mathbf{K} kovariančna matrika učnih vhodov.

Metoda največje podobnosti (ang. maximum likelihood method – ML) da najbolj verjetno oceno hiperparametrov, ki so podlaga za napovedovanje naslednjih izhodov. Če je optimizacija realizirana z metodo konjugiranih gradientov (ali katerokoli drugo gradientno metodo), je potrebno izračunati odvode po vseh parametrih:

$$\frac{\partial \mathcal{L}(\boldsymbol{\Theta})}{\partial \Theta_i} = -\frac{1}{2} \text{trace} \left(\mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \Theta_i} \right) + \frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \Theta_i} \mathbf{K}^{-1} \mathbf{y} \quad (2.7)$$

Ob vsakem koraku optimizacije se zahteva izračun inverza kovariančne matrike \mathbf{K} , kar lahko postane računsko zelo potratno za velike N .

Druga alternativa je, da iz a-priori znanja, ki ga vključimo, izračunamo porazdelitev z numeričnim integriranjem (MCMC metode, glej [12]) po hiperparametrih.

2.5 Napovedovanje

Pogojni del enačbe (2.4) določi porazdelitev izhoda y^* , $p(y^*|\mathbf{y}, \mathbf{X}, \mathbf{x}^*) = \frac{p(\mathbf{y}, y^*)}{p(\mathbf{y}|\mathbf{X})}$. Dobimo normalno porazdelitev s srednjo vrednostjo in varianco kot [10]:

$$\mu(\mathbf{x}^*) = \mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1} \mathbf{y} \quad (2.8)$$

$$\sigma^2(\mathbf{x}^*) = k(\mathbf{x}^*) - \mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}^*) + v_0 \quad (2.9)$$

kjer je $\mathbf{k}(\mathbf{x}^*) = [C(\mathbf{x}^1, \mathbf{x}^*) \dots C(\mathbf{x}^N, \mathbf{x}^*)]$ $N \times 1$ kovariančni vektor med testnim in učnimi vhodi in je $k(\mathbf{x}^*) = C(\mathbf{x}^*, \mathbf{x}^*)$ avtokovarianca testnega vhoda.

Na vektor $\mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1}$ lahko gledamo kot na vektor uteži, ki določa kako bodo uteženi izhodi \mathbf{y} . Če se nov vhod precej razlikuje od učnih, bo izraz $\mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}^*)$ majhen in s tem varianca $\sigma^2(\mathbf{x}^*)$ velika.

Primerjavo z nekaterimi drugimi metodami za regresijo lahko strnemo v naslednjem odstavku. Metode z znano strukturo (npr. metoda najmanjših kvadratov) aproksimirajo neznani sistem z vnaprej izbranimi funkcijami, ki opisujejo strukturo in jim optimiziramo parametre. Regresijske metode kot so umetne nevronske mreže (primerjavo z GP modeli najdemo v [9, 13, 18]) ali mehki sistemi aproksimirajo neznani sistem s primerno uteženimi baznimi funkcijami. Identifikacija z Gaussovimi procesi pa deluje na principu združevanja in glajenja informacije, ki jo nosijo pošumljene meritve. Metoda napoveduje izhod na podlagi razdalje med testnim in učnimi primeri.

2.6 Primer uporabe Gaussovih procesov za identifikacijo statičnega nelinearnega sistema

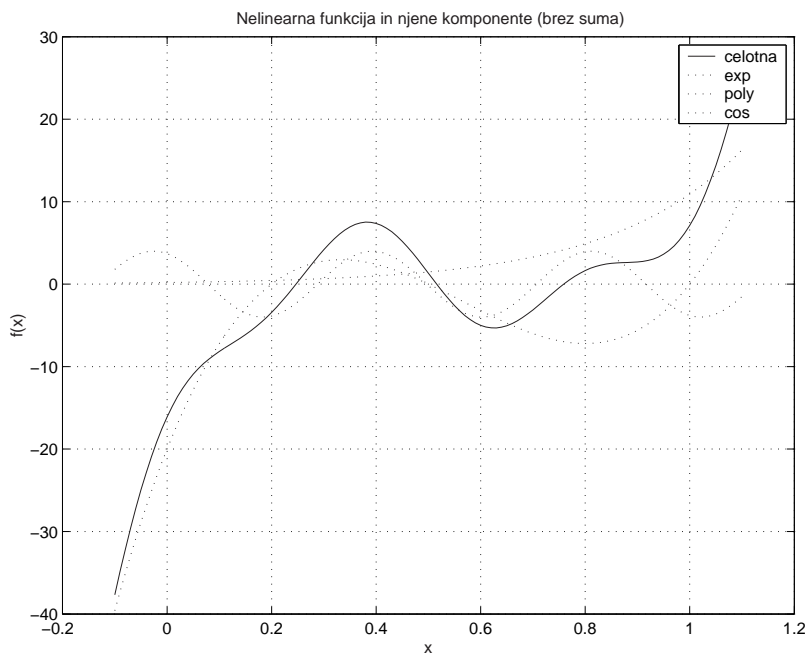
Ilustrirajmo uporabo Gaussovih procesov na primeru identifikacije statičnega nelinearnega sistema.

Želimo identificirati nelinearno funkcijo

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0 + b_0 \exp^{b_1x} + c_0 \cos(c_1x + c_2) \quad (2.10)$$

na intervalu $x \in [0, 1.1]$. Funkcija je kombinacija polinoma tretjega reda, eksponentne ter kosinusne funkcije s parametri $a_3 = 200$, $a_2 = -340$, $a_1 = 160$, $a_0 = -20$, $b_0 = 0.2$, $b_1 = 4$, $c_0 = 4$, $c_1 = 15$, $c_2 = \pi/8$.

Nelinearna funkcija in njeni sestavni deli so prikazani na sliki 2.1. Izhodu funkcije je bil prištet šum s srednjo vrednostjo 0 in standardno deviacijo $\sigma = 1$. Učne točke na intervalu $[0, 1.1]$ so porazdeljene neenakomerno, da lahko preučimo vpliv njihove porazdelitve na identifikacijo.



Slika 2.1: Nelinearna funkcija

Za kovariančno funkcijo je bila izbrana kar običajna Gaussova kovariacijska funkcija, kjer se elementi kovariančne matrice $K_{ij} = C(\mathbf{x}_i, \mathbf{x}_j)$ računajo kot v enačbi (2.1).

Ker je nelinearni sistem določen samo z eno neodvisno spremenljivko, se računanje elementov kovariančne matrike K_{ij} poenostavi v:

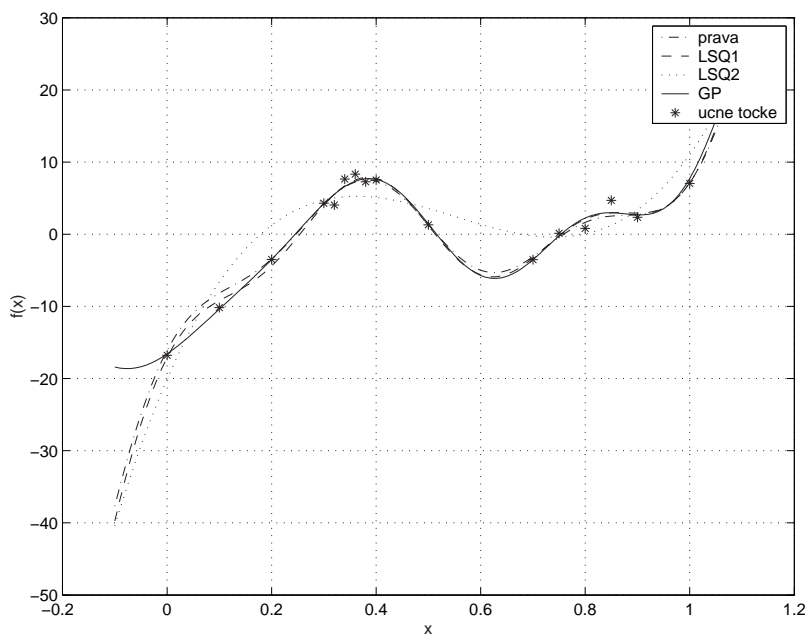
$$C(x_i, x_j) = v \exp \left[-\frac{1}{2} w (x_i - x_j)^2 \right] + \delta v_0 \quad (2.11)$$

Z optimizacijo¹ določimo tri hiperparametre: v , w ter v_0 , ki dobijo vrednosti $w = 26.6357$, $v = 463.1330$ in $v_0 = 0.3027$.

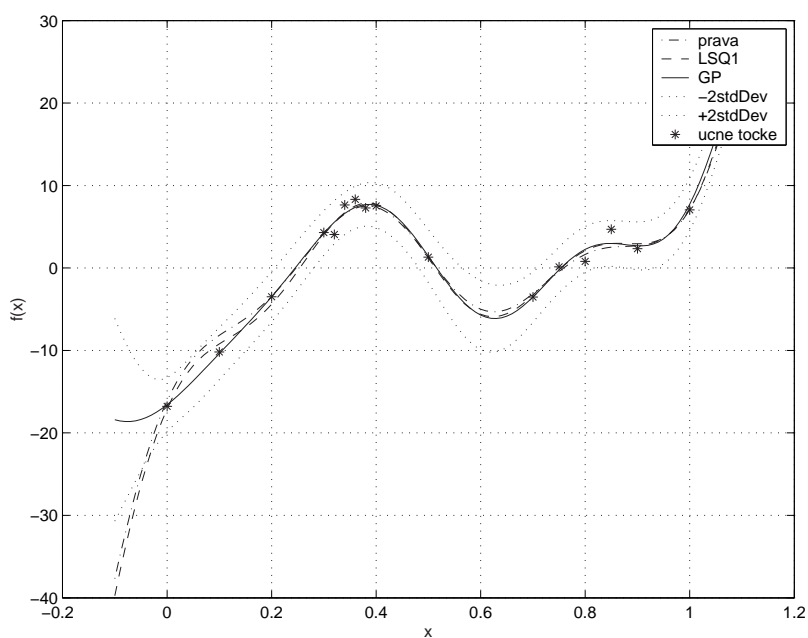
Primerjava rezultatov identifikacije z GPi in identifikacije z metodo najmanjših kvadratov (LSQ) je prikazana na slikah 2.2 in 2.3. Kot izhod iz GPa je prikazana srednja vrednost izhodne porazdelitve. Učne točke so označene z zvezdico. Slabost metode najmanjših kvadratov je, da daje precej slabše rezultate, če ne poznamo točne strukture identificirane funkcije. Z LSQ1 je označena metoda LSQ, kjer je struktura identificirane funkcije znana in enaka originalni funkciji, z LSQ2 pa metoda LSQ, kjer predpostavimo, da je neznana funkcija polinom tretjega reda. Na sliki 2.3 vidimo primerjavo identifikacije z GPi in LSQ metodo pri poznavanju strukture identificirane funkcije. Metoda z GPi da malenkost slabše, a primerljive rezultate. Na sliki 2.2 je prikazan tudi rezultat identifikacije z metodo LSQ2, ko struktura identificirane funkcije ni popolnoma znana. Ta metoda LSQ precej slabše opravi svoje delo.

Slika 2.4 prikazuje absolutne vrednosti napake metod LSQ1 in GP ter absolutne vrednosti napak učnih točk. Obe metodi imata največjo napako na področju, kjer je malo učnih točk in so le te močno pošumljene. Prav tako sta na sliki 2.4 predstavljeni standardni deviaciji obeh metod. Varianca metode z GPi je sestavljena iz dveh delov – en del predstavlja varianco odziva v_0 zaradi šuma, drug del pa varianco metode. Metoda LSQ daje samo varianco metode, ki je po celotnem vhodnem prostoru enaka. Na sliki 2.4 opazimo zmanjšanje negotovosti na intervalu $[0.3, 0.4]$, kjer je večja gostota učnih točk. Prav tako opazimo, da je funkcija povsod znotraj intervala 95% verjetnosti. Zanimiv je tudi rezultat identifikacije na območju $x < 0$ in intervalu $[0.5, 0.7]$, kjer ni učnih točk. Napaka predikcije GP je večja, a nas nanjo opozori velika varianca.

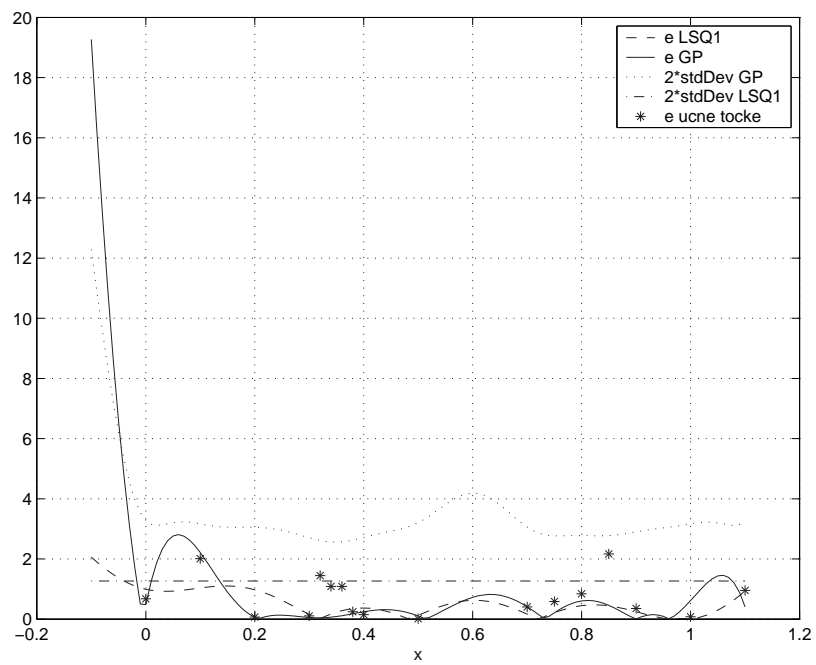
¹uporabljene rutine za programski paket Matlab C.E. Rasmussen-a, <http://www.gatsby.ucl.ac.uk/~edward/>



Slika 2.2: Primerjava identifikacije z metodo najmanjših kvadratov (LSQ) in z GPi



Slika 2.3: Primerjava identifikacije z metodo LSQ z znano strukturo funkcije in identifikacije z GPi



Slika 2.4: Primerjava napake predikcije metod LSQ in GP

3. Uporaba Gaussovih procesov za identifikacijo nelinearnih dinamičnih sistemov

3.1 Identifikacija dinamičnih sistemov

Čeprav se Gaussovi procesi že nekaj let uporabljajo za identifikacijo statičnih nelinearnosti, pa je identifikacija dinamičnih sistemov z Gaussovimi procesi (GP) novejšo področje. Šele v zadnjih letih je bilo objavljenih nekaj del s to tematiko, npr. [3, 6, 10, 9].

Identifikacija z Gaussovimi procesi deluje na principu statične preslikave vhodov na izhod preko kovariančne matrike \mathbf{K} (enačba 2.1). Dinamični sistemi so za razliko od statičnih sistemov poleg trenutne vrednosti vhodnega signala v splošnem odvisni tudi od preteklih vrednosti vhodov in izhodov.

Naj naš sistem predstavlja funkcija

$$y(k) = f(y(k-1), \dots, y(k-n), u(k-1), \dots, u(k-n)) \quad (3.1)$$

Tu se pojavita dve možnosti za ponazoritev dinamičnega sistema z GPi. Prva je, da kot vhod uporabimo napovedano srednjo vrednost signala, kar bo vsebina razdelka 3.3. To je tako imenovana "naivna" metoda. Druga možnost je, da kot vhod v GP nastopa celotna izhodna porazdelitev (ki je po predpostavki Gaussova). Teji metodi rečemo "eksaktna", opisana pa je v razdelku 3.4. V razdelku 3.2 pa bomo predstavili dva dinamična sistema, s katerima bomo prikazali identifikacijo.

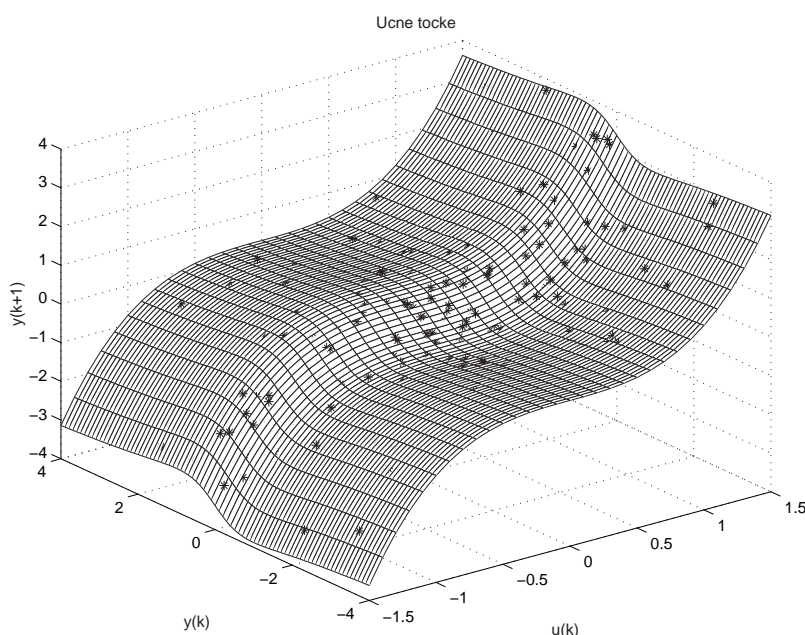
3.2 Primera identifikacije nelinearnih dinamičnih sistemov

Za ilustracijo delovanja metode GP za identifikacijo smo uporabili dva nelinearna dinamična sistema.

Prvi izbran nelinearni dinamični sistem, je bil uporabljen tudi kot testni primer za identifikacijo z umetnimi nevronskimi mrežami v delu [17]. Gre za nelinearni diskretni sistem prvega reda, katerega enačba se glasi:

$$y(k+1) = \frac{y(k)}{1+y^2(k)} + u^3(k) \quad (3.2)$$

Na sliki 3.1 je prikazano obnašanje sistema na najbolj zanimivem območju okoli izhodišča $y(k) = u(k) = 0$. Bližnja okolica te točke je bila tudi izbrana za identifikacijo (področje identifikacije je omejeno z $\max |u(k)| = 1.5$ in posledično $\max |y(k)| \simeq 3.9$).



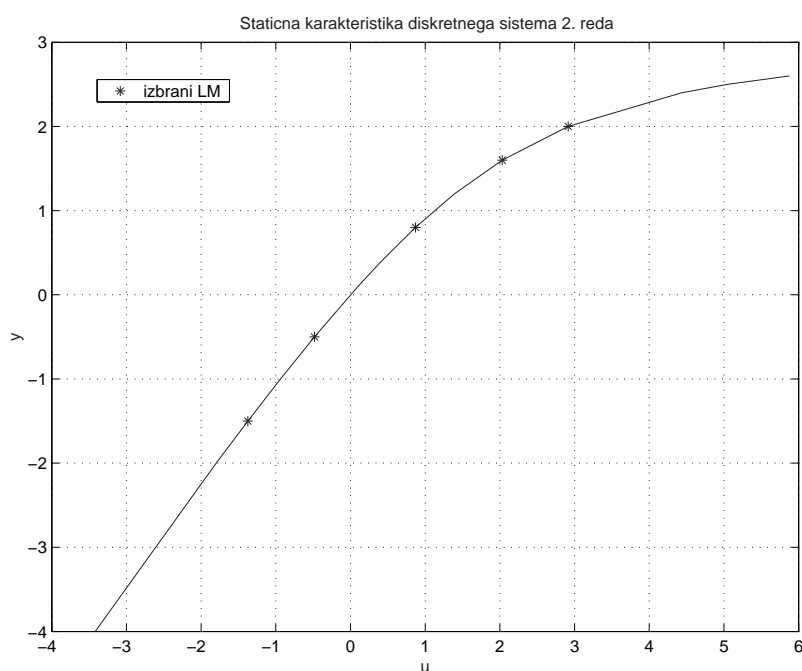
Slika 3.1: Nelinearni dinamični sistem prvega reda in učne točke

Drugi izbrani nelinearni dinamični sistem je bil prav tako diskreten, a drugega reda, uporabljen v [15]. Njegovo obnašanje predstavlja diferenčna enačba (3.3):

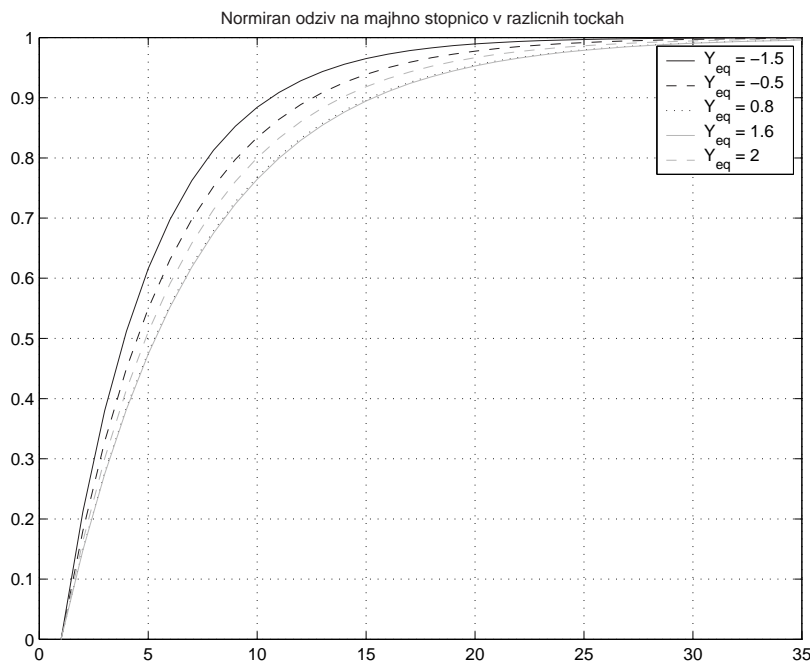
$$y(k) = 0.893y(k-1) + 0.0371y^2(k-1) - 0.05y(k-2) - 0.05u(k-1) + 0.0157u(k-1) \quad (3.3)$$

Sistem je drugega reda in ima tri regresorje, zato ga ne moremo predstaviti v tridimenzionalnem prostoru kot smo predstavili sistem prvega reda.

Statično karakteristiko sistema (3.3) na modeliranem področju si lahko ogledamo na sliki 3.2. Izbrano je bilo področje delovanja med $u_{min} = -2$ in $u_{max} = 4$. Na sliki 3.3 so prikazani normirani odzivi na (majhno) stopnico v petih točkah ($y_1 = -1.5, y_2 = -0.5, y_3 = 0.8, y_4 = 1.6, y_5 = 2.0$).



Slika 3.2: Statična karakteristika nelinearnega dinamičnega sistema drugega reda



Slika 3.3: Normiran odziv na stopnico v različnih točkah nelinearnega dinamičnega sistema drugega reda

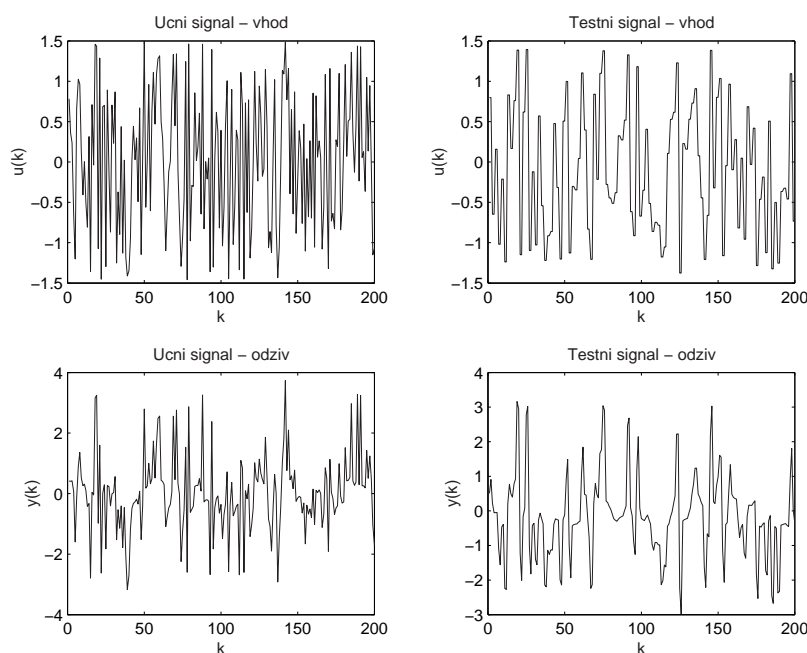
3.3 Predikcija pri napovedani srednji vrednosti kot vhodu v GP model

3.3.1 Dinamični sistem prvega reda

Učne signale (tudi identifikacijske) in testne signale (tudi signale za vrednotenje) smo dobili tako, da smo tvorili različne naključne signale u in z njimi vzbujali sistem. Tako učni kot testni signal u (slika 3.4) sta bila omejena na interval $[-1.5 \ 1.5]$. Identificirali smo model prvega reda z dvema regresorjema ($d = 2$): vrednostjo vzbujevalnega signala $u(k)$ in srednjo vrednostjo predikcije izhoda, ki jo bomo označili z $\hat{y}(k)$. Regresorja sestavljata vhodni vektor, želeni izhod (ang. target) pa je izhod iz sistema $y(k + 1)$, ki mu je dodan šum z varianco $\sigma^2 \simeq 0.04$.

Uporabljena je bila standardna kovariančna funkcija (2.1) in optimizirali smo štiri hiperparametre: v, w_1, w_2 ter varianco šuma v_0 .

Za učenje je bilo uporabljenih dvesto vhodno-izhodnih parov podatkov. Število je bilo izbrano kot kompromis med natančnostjo in trajanjem optimizacije (čas za računanje inverzne matrike raste s tretjo potenco ranga [2], inverz

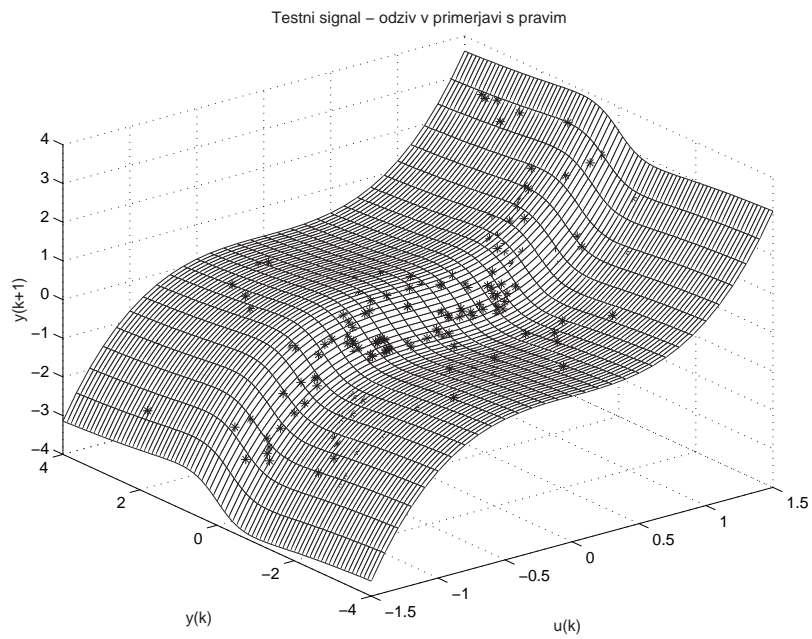


Slika 3.4: Signala za učenje in vrednotenje – dinamični sistem prvega reda

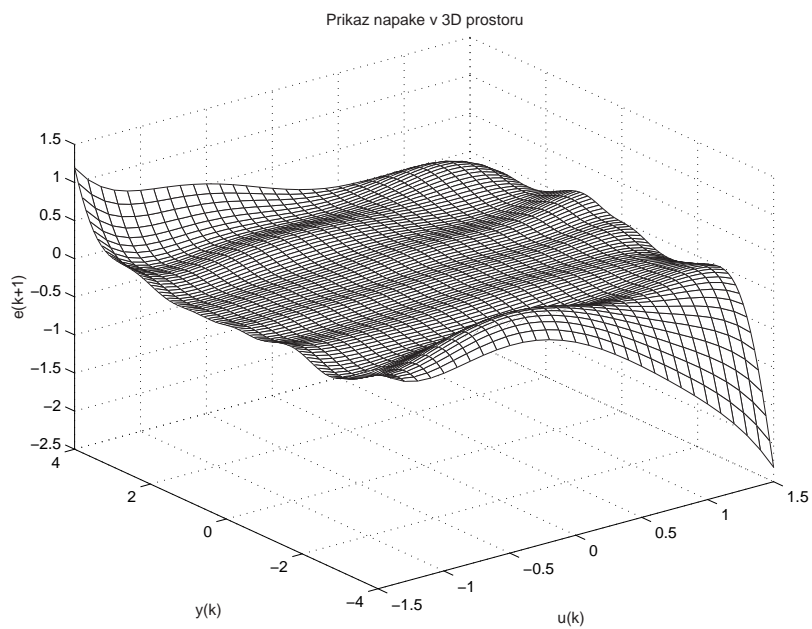
matrike pa se računa v vsakem koraku optimizacije). Optimizacija je imela dvesto korakov. Tudi testni signal je imel dvesto vzorcev, a drug takt spremembe amplitude signala kot učni signal.

Na sliki 3.5 vidimo srednjo vrednost izhoda iz GPa, na sliki 3.6 pa njegovo napako. Varianca je predstavljena na sliki 3.7. Na teh slikah so predstavljeni rezultati enokoračne predikcije. Bolj pomembno vrednotenje modela pa je simulacija, kjer srednje vrednosti izhoda (ne pozabimo, da je izhod iz GPa normalna porazdelitev) s povratno zanko peljemo nazaj na vhod. Rezultati simulacije so prikazani na sliki 3.8. Ti so dobri za področje, kjer so obstajali učni podatki. Ko pa se pomaknemo izven tega področja, napaka hitro naraste. Dobra lastnost Gaussovih procesov je, da smo na to opozorjeni s povečanjem variance. Oboje je lepo vidno na robnih območjih na slikah 3.5, 3.6 in 3.7.

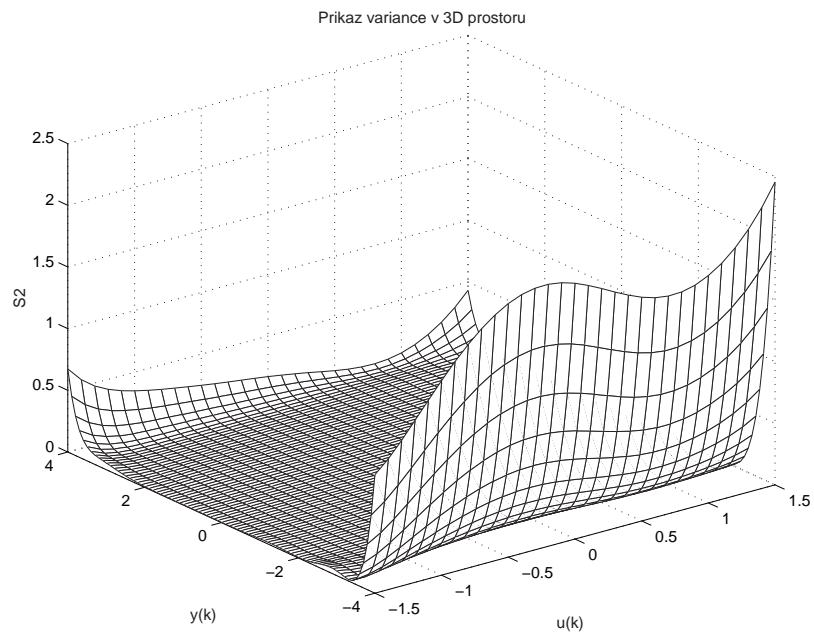
Statistična pokazatelja kvalitete identifikacije, avtokorelacijo napake $\Phi_{ee}(k)$ in križno korelacijo $\Phi_{ue}(k)$ med vhodnim signalom $u(k)$ in napako $e(k+1)$, ki kažeta na zadovoljive rezultate, vidimo na slikah 3.9 in 3.10. Za ocenitev odzivov bomo uporabljali cenilki za povprečno vrednost kvadrata napake (ang. mean squared error)



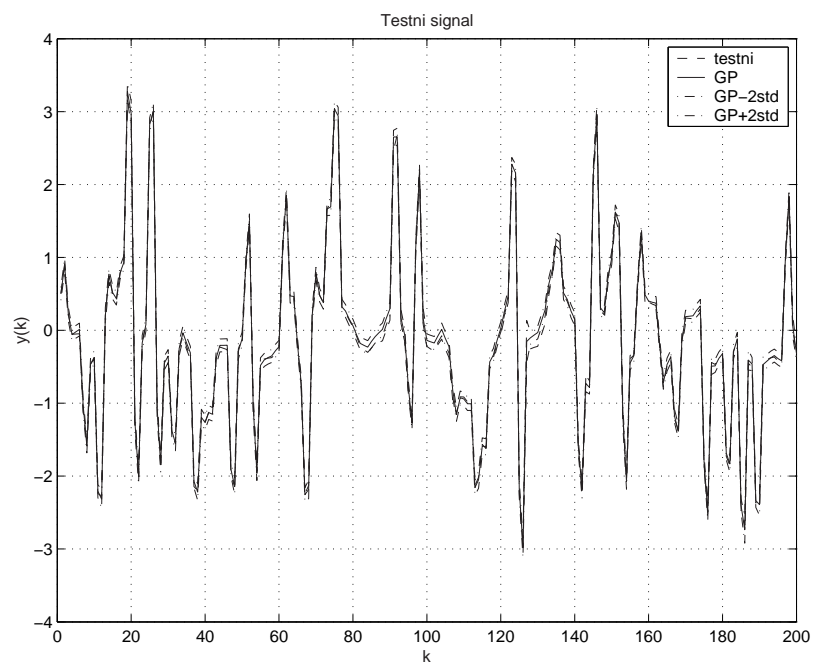
Slika 3.5: Napovedi GP modela in nelinearnost originalnega sistema



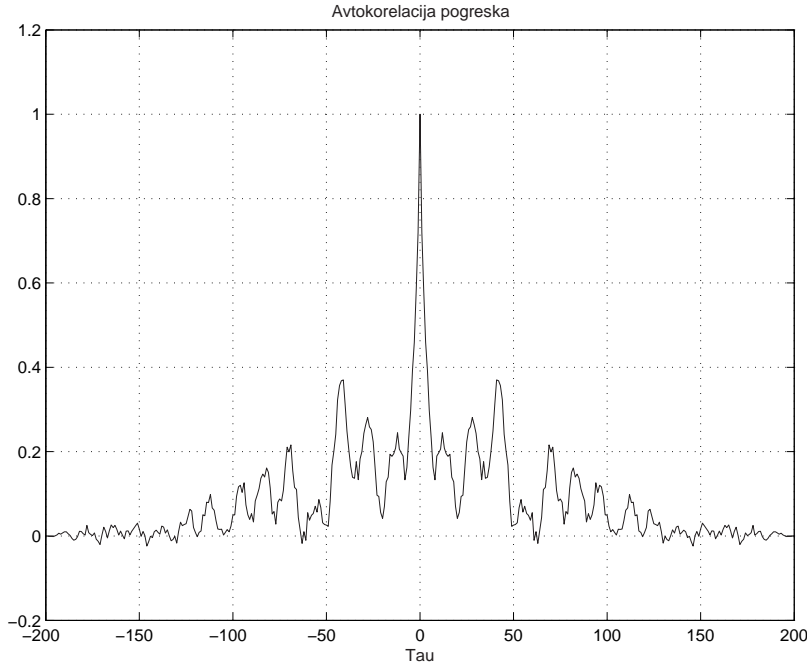
Slika 3.6: Napaka GP



Slika 3.7: Varianca GP



Slika 3.8: Simulacija na testnem signalu

Slika 3.9: Avtokorelacija napake $\Phi_{ee}(\tau)$

$$SE = \frac{1}{N} \sum_{i=1}^N e_i^2 \quad (3.4)$$

in logaritem gostote napake odziva (ang. log-predictive density error, [21]), ki se pogosto uporablja v povezavi z metodami največje podobnosti:

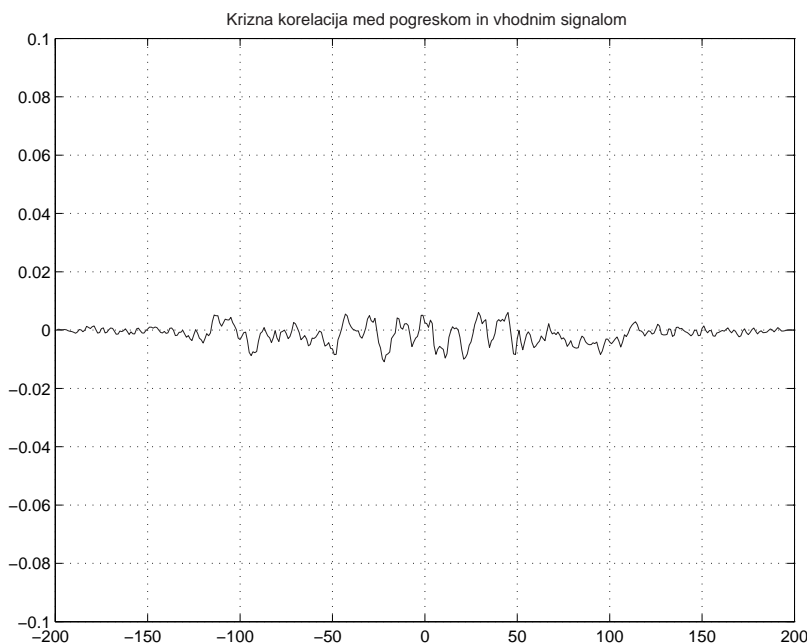
$$LD = \frac{1}{2N} \sum_{i=1}^N \left(\log(2\pi) + \log(\text{var}_i) + \frac{e_i^2}{\text{var}_i} \right) \quad (3.5)$$

kjer je $e_i = \hat{y}_i - y_i$ napaka predikcije GP modela v i -tem koraku in var_i napovedana varianca GP modela v i -tem koraku.

Povprečna napaka odziva GP modela je $SE = 0.0094$, logaritem gostote napake odziva pa $LD = -1.7555$.

3.3.2 Dinamični sistem drugega reda

Za vsako uspešno identifikacijo so potrebni dobri učni podatki. Naj je metoda še tako dobra, brez pravilno izbranih učnih podatkov ne more dobro identificirati obravnavanega sistema. Delovanje GP modela temelji na razdalji med testnim



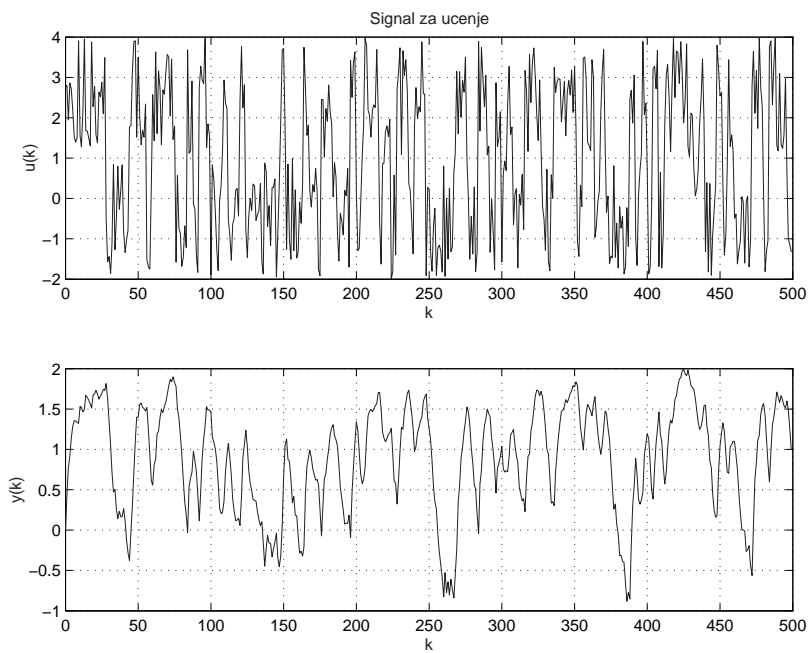
Slika 3.10: Križna korelacija med vzbujevalnim vhodnim signalom in napako $\Phi_{ue}(\tau)$

podatkom in učnimi podatki. Bližje (velja za standardno kovariančno funkcijo (2.1)) kot je testni podatek kateremu izmed učnih, večja je njuna povezanost (korelacija) in večje je matematično upanje, da vrednost napovedi pri določenem testnem podatku ne bo zelo oddaljena od vrednosti napovedi pri tem učnem podatku. Zato je pomembno izbrati take učne točke, da čim bolj pokrijemo celotno želeno področje identifikacije. Sistemi prvega reda imajo samo dva regresorja, tako da lahko vhodno informacijo skupaj z izhodom predstavimo v tridimenzionalnem prostoru in si olajšamo izbiro učnih podatkov. Pri modelih višjega reda z več kot dvema vhodoma in enim izhodom predstavitev vhodne informacije in njen vpliv na izhod nista več tako enostavna, saj si ne moremo pomagati s prostorsko predstavitvijo.

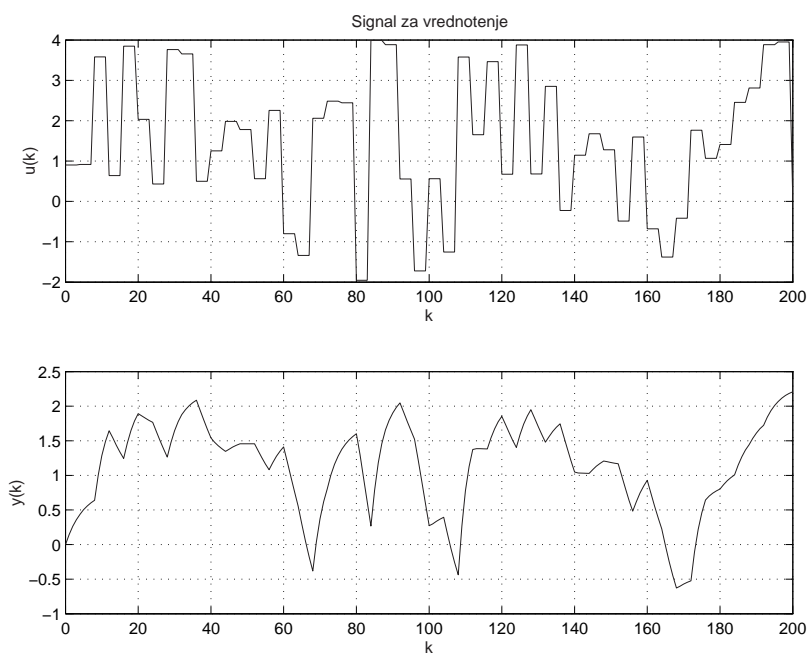
Signala za učenje in vrednotenje sta bila tvorjena enako: kot vhod v proces je nastopal naključni vhodni signal $u(k)$, omejen med vrednosti $u_{min} = -2$ in $u_{max} = 4$. Prikazana sta na slikah 3.11 in 3.12.

Želenemu izhodu učnega signala je bil prištet beli šum z varianco $\sigma^2 = 4 \cdot 10^{-4}$.

Za učenje je bilo uporabljenih 83 statistično neodvisno izbranih učnih točk iz



Slika 3.11: Signal za učenje dinamičnega sistema drugega reda

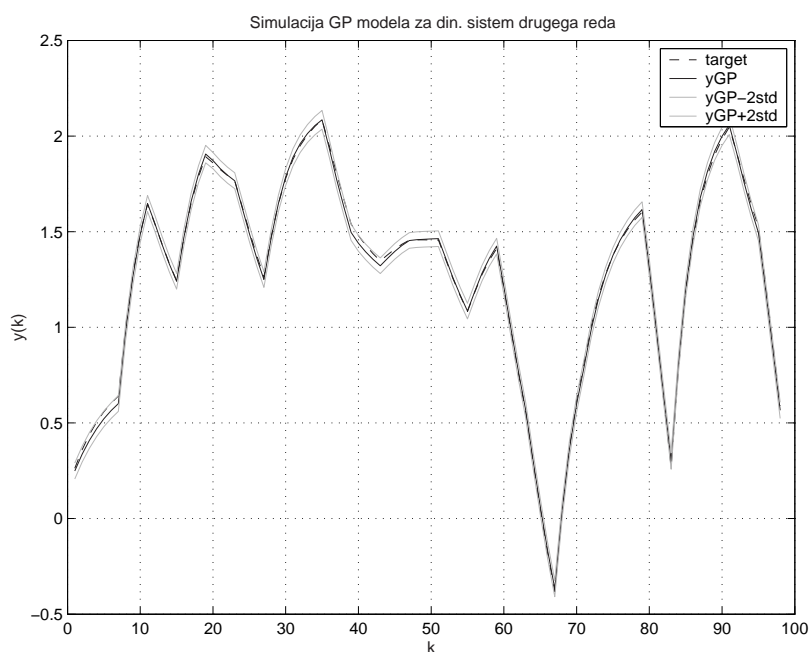


Slika 3.12: Signal za vrednotenje dinamičnega sistema drugega reda

učnega signala. Rezultat simulacije identificiranega GP modela je prikazan na sliki 3.13¹. Predikcija vseskozi ostane znotraj meja zaupanja $\pm 2\sigma$. Na sliki 3.14 je prikazana absolutna napaka simulacije $|e(k)| = |y_{GP}(k) - y(k)|$ skupaj z napovedano dvakratno standardno deviacijo 2σ .

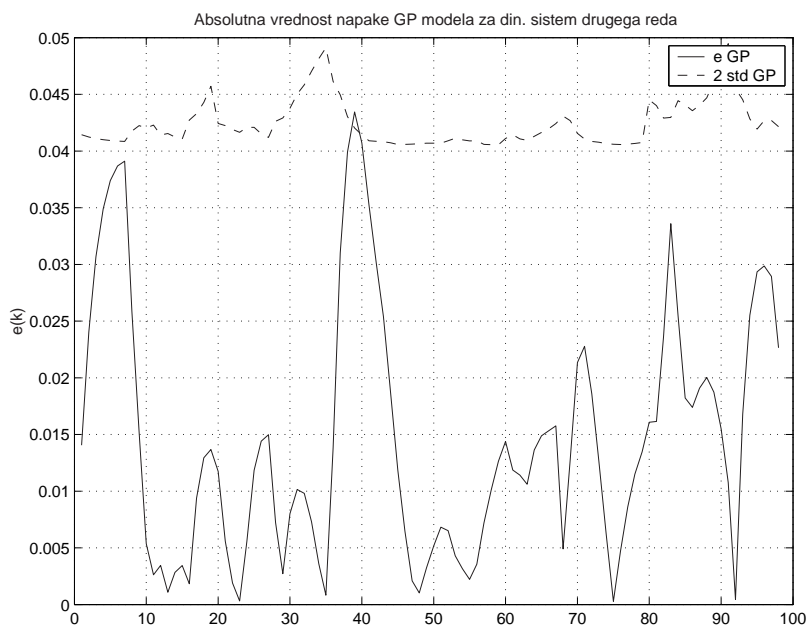
Sliki 3.15 in 3.16 prikazujeta avtokorelacijo napake simulacije $e(k)$ in križno korelacijo med napako identifikacije $e(k)$ ter vhomom v sistem $u(k)$. Vrednosti cenilk odziva GP modela sta $SE = 3.3E - 4$ in $LD = -2.56$.

Rezultati identifikacije nelinearnega dinamičnega sistema drugega reda zado-
stujejo našemu namenu, to je prikazati identifikacijo sistema z modelom GP.

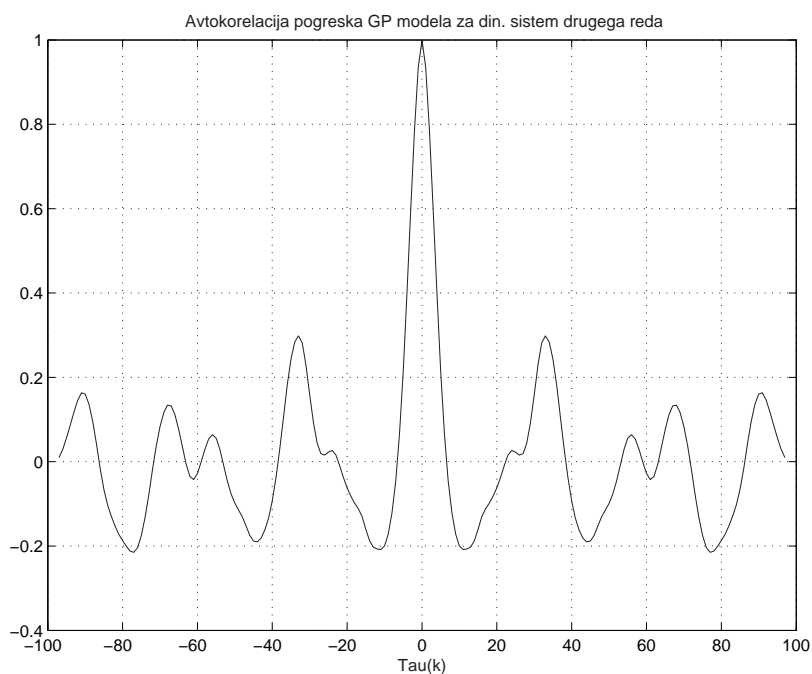


Slika 3.13: Simulacija GP modela za nelinearni dinamični sistem drugega reda

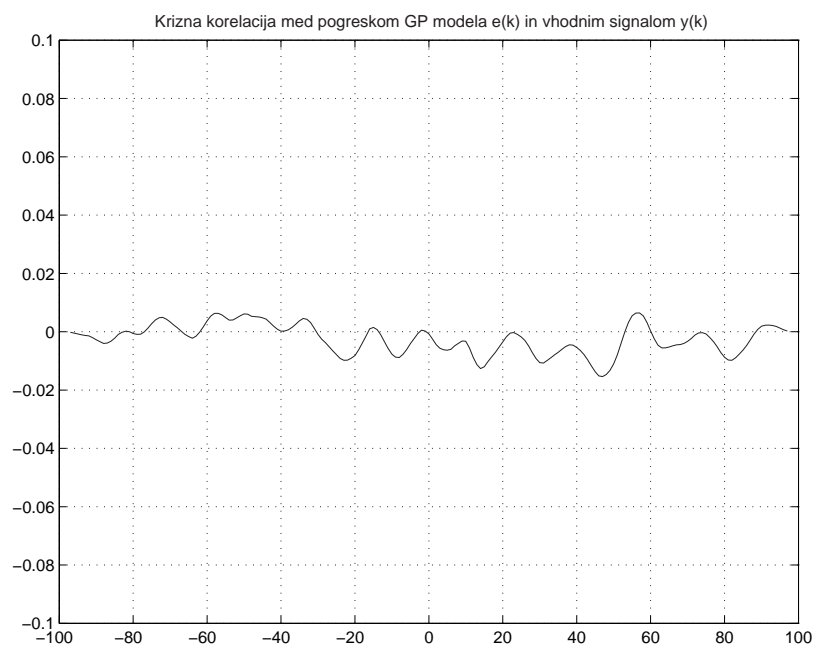
¹Ker pozneje identificiramo še en sistem drugega reda, je ponekod za večjo nazornost v naslov slike dodana beseda umetni za trenutno obravnavani in naprava za kasneje predstavljeni sistem.



Slika 3.14: Absolutna napaka $|e(k)|$ simulacije GP modela, nelinearni dinamični sistem drugega reda



Slika 3.15: Avtokorelacija napake predikcije GP modela Φ_{ee} , nelinearni dinamični sistem drugega reda



Slika 3.16: Križna korelacija vhoda in napake predikcije GP modela Φ_{ue} , nelinearni dinamični sistem drugega reda

3.4 Predikcija pri porazdelitvi kot vhodu v GP model

V prejšnjem razdelku smo pokazali, kako najenostavneje simuliramo dinamični sistem (iterativna enokoračna predikcija). Kot novi vhod smo uporabili pretekle srednje vrednosti izhoda. Kot smo že povedali, imenujemo tako predikcijo "naivna". Njena slabost je, da zavržemo informacijo o porazdelitvi izhoda, s tem pa tudi informacijo o zaupanju vrednosti, ki jo uporabimo kot nov vhod. V tem poglavju bomo opisali, kako koristno uporabiti tudi podatek o negotovosti srednje vrednosti izhoda – varianci.

Še vedno nas zanima napoved izhodnega signala \hat{y} , namesto srednjih vrednosti izhoda pa sedaj kot vhod nastopajo napovedane porazdelitve izhoda. Tak GP model ne samo da opisuje dinamične lastnosti nelinearnega sistema, ampak da tudi informacijo o stopnji zaupanja v predikcijo. Za napovedovanje novega izhoda imamo dve možnosti [3, 4]:

- numerično integriranje po vhodni porazdelitvi z Monte Carlo metodo
- analitično rešitev, pri čemer uvedemo določene poenostavitve [3, 4]. Utemeljitev teh poenostavitev je podana in ilustrirana v [5, 4]. Predikcijo v tem pristopu imenujemo "eksaktna" predikcija.

Ker so izpeljave izrazov, ki se nahajajo v [3, 4] obsežne, se bomo omejili na glavne korake v izpeljavi.

Izhajajoč iz [3, 4], kjer je predpostavljeno, da je izhod normalno porazdeljen s srednjo vrednostjo m in varianco v tudi za naključno spremenljivko $\mathbf{x}^* = \int p(\mathbf{x})d\mathbf{x}$ na vhodu, za novo oceno izhoda iz modela $\hat{y}(k)$ dobimo $\hat{y}(k) \sim \mathcal{N}(m(\mathbf{x}(k), v(\mathbf{x}(k)) + v_0)$, kjer sta srednja vrednost $m(\cdot)$ in varianca $v(\cdot)$ izračunana z enačbama (3.6) in (3.7).

$$m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_{\mathbf{x}}) = E_{\mathbf{x}^*}[\mu(\mathbf{x}^*)] \quad (3.6)$$

$$v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_{\mathbf{x}}) = E_{\mathbf{x}^*}[\sigma^2(\mathbf{x}^*)] + E_{\mathbf{x}^*}[\mu(\mathbf{x}^*)^2] - (E_{\mathbf{x}^*}[\mu(\mathbf{x}^*)])^2 \quad (3.7)$$

Vektor $\boldsymbol{\nu}$ sestavljajo srednje vrednosti regresorjev, matriko $\boldsymbol{\Sigma}_{\mathbf{x}}$ pa vrednosti križnih kovarianc med posameznimi vhodi.

Posamezni elementi enačb (3.6) in (3.19) so:

$$E_{\mathbf{x}^*}[\mu(\mathbf{x}^*)] = \sum_{i=1}^N \beta_i \int C(\mathbf{x}, \mathbf{x}_i) p(\mathbf{x}) d\mathbf{x} \quad (3.8)$$

$$E_{\mathbf{x}^*}[\sigma^2(\mathbf{x}^*)] = C(\boldsymbol{\nu}, \boldsymbol{\nu}) - \sum_{i,j=1}^N K_{ij}^{-1} \int C(\mathbf{x}, \mathbf{x}_i) C(\mathbf{x}, \mathbf{x}_j) p(\mathbf{x}) d\mathbf{x} \quad (3.9)$$

$$E_{\mathbf{x}^*}[\mu(\mathbf{x}^*)^2] = \beta_i \beta_j \sum_{i,j=1}^N K_{ij}^{-1} \int C(\mathbf{x}, \mathbf{x}_i) C(\mathbf{x}, \mathbf{x}_j) p(\mathbf{x}) d\mathbf{x} \quad (3.10)$$

$$(3.11)$$

kjer je β_i i -ti element vektorja $\boldsymbol{\beta} = \mathbf{K}^{-1}\mathbf{y}$, $C(\boldsymbol{\nu}, \boldsymbol{\nu})$ vrednost avtokovariančne funkcije (2.1) za napovedane srednje vrednosti in $C(\mathbf{x}, \mathbf{x}_i)$ vrednost kovariančne funkcije (2.1) med vhom \mathbf{x} in učnim vhom \mathbf{x}_i .

Integrala v zgornjih enačbah

$$I_1 = \int C(\mathbf{x}, \mathbf{x}_i) p(\mathbf{x}) d\mathbf{x} \quad (3.12)$$

$$I_2 = \int C(\mathbf{x}, \mathbf{x}_i) C(\mathbf{x}, \mathbf{x}_j) p(\mathbf{x}) d\mathbf{x} \quad (3.13)$$

sta analitično rešljiva le za malo oblik kovariančnih funkcij $C(\mathbf{x}_i, \mathbf{x}_j)$. Za nas je pomembno, da za kovariančno funkcijo z elementi oblike (2.1) velja:

$$I_1 = \int N_x(\mathbf{x}_i, \mathbf{W}) N_x(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) d\mathbf{x} = \quad (3.14)$$

$$= N_u(\mathbf{x}_i, \mathbf{W} + \boldsymbol{\Sigma}_x) \quad (3.15)$$

$$I_2 = \int N_x(\mathbf{x}_i, \mathbf{W}) N_x(\mathbf{x}_j, \mathbf{W}) N_x(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) d\mathbf{x} = \quad (3.16)$$

$$= N_{x_i}(\mathbf{x}_j, 2\mathbf{W}) N_u\left(\frac{\mathbf{x}_i + \mathbf{x}_j}{2}, \frac{\mathbf{W}}{2} + \boldsymbol{\Sigma}_x\right) \quad (3.17)$$

kjer je z N označena predpostavljena normalna porazdelitev \mathcal{N} in je \mathbf{W} diagonalna matrika z inverzi hiperparametrov w_1, \dots, w_d na diagonali (2.3).

Pri predpostavki, da ima izhod GP normalno porazdelitev, je bilo pokazano [8], da srednjo vrednost $m(\cdot)$ in varianco $v(\cdot)$ izračunamo kot:

$$m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_{\mathbf{x}}) = \sum_{i=1}^N \beta_i C_{\text{mod}_1}(\mathbf{x}^*, \mathbf{x}_i) \quad (3.18)$$

$$v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_{\mathbf{x}}) = v + \sum_{i,j=1}^N (\beta_i \beta_j - K_{ij}^{-1}) C_{\text{mod}_2}(\mathbf{x}_i, \mathbf{x}_j) C_{\text{mod}_3}(\mathbf{x}^*, \mathbf{x}_b) - m^2(\boldsymbol{\nu}, \boldsymbol{\Sigma}_{\mathbf{x}}) \quad (3.19)$$

kjer je \mathbf{x}_i i -ti učni vhodni regresor in so

$$C_{\text{mod}_1}(\mathbf{x}^*, \mathbf{x}_i) = v \left| \mathbf{I} + \mathbf{W}^{-1} \boldsymbol{\Sigma}_x \right|^{-1/2} \exp \left[-\frac{1}{2} (\mathbf{x}^* - \mathbf{x}_i)^T (\mathbf{W} + \boldsymbol{\Sigma}_x)^{-1} (\mathbf{x}^* - \mathbf{x}_i) \right] \quad (3.20)$$

$$C_{\text{mod}_2}(\mathbf{x}_i, \mathbf{x}_j) = v_0 2^{-\frac{D}{2}} \exp \left[-\frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^T \left(\frac{\mathbf{W}}{2} \right)^{-1} (\mathbf{x}_i - \mathbf{x}_j) \right] \quad (3.21)$$

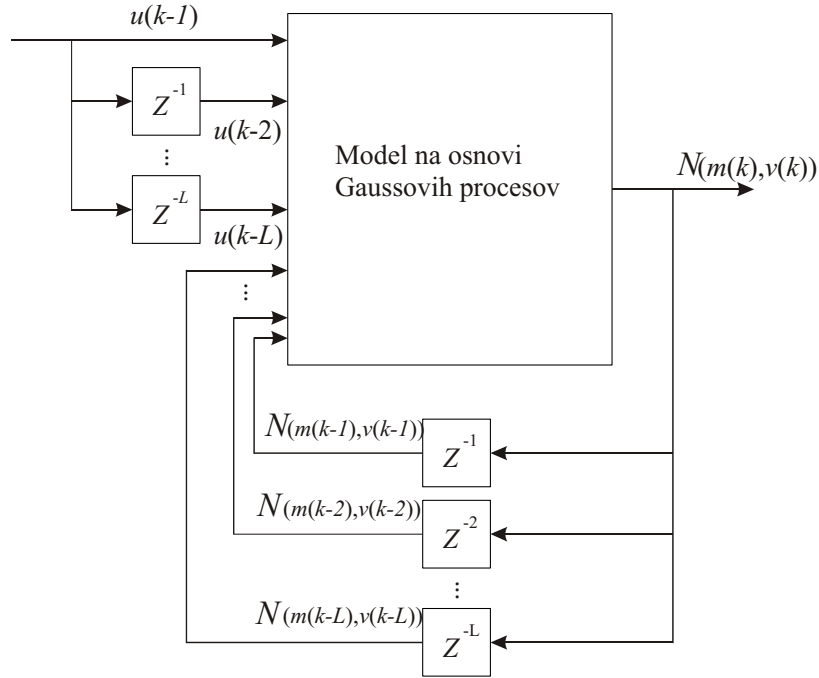
$$C_{\text{mod}_3}(\mathbf{x}^*, \mathbf{x}_b) = v_0 \left| \frac{1}{2} \mathbf{I} + \mathbf{W}^{-1} \boldsymbol{\Sigma}_x \right|^{-1/2} \exp \left[-\frac{1}{2} (\mathbf{x}^* - \mathbf{x}_b)^T \left(\frac{\mathbf{W}}{2} + \boldsymbol{\Sigma}_x \right)^{-1} (\mathbf{x}^* - \mathbf{x}_b) \right] \quad (3.22)$$

ter β_i elementi vektorja $\boldsymbol{\beta} = \mathbf{K}^{-1} \mathbf{y}$ in $\mathbf{x}_b = \frac{\mathbf{x}_i + \mathbf{x}_j}{2}$.

Ideja simulacije GP modela je prikazana na sliki 3.17 in povzeta po [3]. Predpostavimo, da poznamo vrednosti izhodov do k -tega koraka. Zanima nas predikcija za n korakov naprej. Velja:

- $k + 1$: $\mathbf{x}_{k+1} = [y(k), \dots, y(k - L + 1)]^T$, kjer smo sedaj z \mathbf{x}_k označili vektor L preteklih srednjih vrednosti izhoda. Ker poznamo vrednost izhodov do k -tega koraka, se srednja vrednost $\mu(\mathbf{x}_{k+1})$ in varianca $\sigma^2(\mathbf{x}_{k+1})$ novega napovedanega izhoda $\hat{y}(k + 1)$ izračunata preprosto z uporabo (3.6) in (3.19).
- $k + 2$: $\mathbf{x}_{k+2} = [\hat{y}(k + 1), y(k), \dots, y(k - L + 2)]^T$. Zdaj kot vhod v model že nastopa predikcija izhoda v prejšnjem koraku simulacije $\hat{y}(k + 1) \sim \mathcal{N}(\mu(\mathbf{x}_{k+1}), \sigma^2(\mathbf{x}_{k+1}))$:

$$\mathbf{x}_{k+2} \sim \mathcal{N} \left(\begin{bmatrix} \mu(\mathbf{x}_{k+1}) \\ y(k) \\ \vdots \\ y(k - L + 2) \end{bmatrix}, \begin{bmatrix} \sigma^2(\mathbf{x}_{k+1}) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & 0 \end{bmatrix} \right)$$



Slika 3.17: Simulacija nelinearnega dinamičnega sistema z GP modelom

Srednja vrednost in varianca novega izhoda $\hat{y}(k+2)$ se ponovno izračunata z uporabo (3.6) in (3.19).

- $k+3$: $\mathbf{x}_{k+3} = [\hat{y}(k+2), \hat{y}(k+1), \dots, y(k-L+3)]^T$, kjer $\hat{y}(k+1) \sim \mathcal{N}(\mu(\mathbf{x}_{k+1}), \sigma^2(\mathbf{x}_{k+1}))$ in $\hat{y}(k+2) \sim \mathcal{N}(m(\mathbf{x}_{k+2}), v(\mathbf{x}_{k+2}))$. Imamo:

$$\mathbf{x}_{k+3} \sim \mathcal{N} \left(\begin{bmatrix} m(\mathbf{x}_{k+2}) \\ \mu(\mathbf{x}_{k+1}) \\ y(k) \\ \vdots \\ y(k-L+3) \end{bmatrix}, \begin{bmatrix} v(\mathbf{x}_{k+2}) & \text{Cov}[y(k+2), y(k+1)] & 0 & \dots & 0 \\ \text{Cov}[y(k+1), y(k+2)] & \sigma^2(\mathbf{x}_{k+1}) & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix} \right)$$

Napoved izhoda $\hat{y}(k+3)$ s srednjo vrednostjo $m(\mathbf{x}_{k+3})$ in varianco $v(\mathbf{x}_{k+3})$ ponovno dobimo z uporabo (3.6) in (3.19).

- $k+n$

V $k+n$ -tem koraku tako kot vhod v model nastopajo napovedi porazdelitve preteklih izhodov \mathbf{x}_{k+n} , dane z vektorjem srednjih vrednosti napovedi $\boldsymbol{\nu}$ in matriko križnih kovarianc $\boldsymbol{\Sigma}_{\mathbf{x}}$:

$\mathbf{x}_{k+n} = [y(k+n-1), y(k+n-2), \dots, y(k+n-L)]^T$, kjer

$$\mathbf{x}_{k+n} \sim \mathcal{N} \left(\begin{bmatrix} m(\mathbf{x}_{k+n-1}) \\ m(\mathbf{x}_{k+n-2}) \\ \dots \\ m(\mathbf{x}_{k+n-L}) \end{bmatrix}, \begin{bmatrix} v(x_{k+n-1}) & \text{Cov}[y(k+n-1), y(k+n-2)] & \dots & \text{Cov}[y(k+n-1), y(k+n-L)] \\ \text{Cov}[y(k+n-2), y(k+n-1)] & v(x_{k+n-2}) & \dots & \text{Cov}[y(k+n-2), y(k+n-L)] \\ \dots & \dots & \dots & \dots \\ \text{Cov}[y(k+n-L), y(k+n-1)] & \text{Cov}[y(k+n-L), y(k+n-2)] & \dots & v(x_{k+n-L}) \end{bmatrix} \right) \quad (3.23)$$

Pripadajoči izhod iz modela s srednjo vrednostjo $m(\mathbf{x}_{k+n})$ in varianco $v(\mathbf{x}_{k+n})$ se ponovno izračuna z uporabo (3.6) in (3.19).

Ob koraku k moramo, da lahko stopimo korak naprej, izračunati križno kovarianco $\text{Cov}(\hat{y}(k+n), \hat{y}(k+n-j))$ med pravkar napovedanim izhodom $\hat{y}(k+n)$ in preostalimi izhodi $\hat{y}(k+n-j)$ do reda sistema L . Križna kovarianca med posameznima napovedanima izhodoma $y(k+n)$ in $y(k+n-j)$ je dana z izrazom [3, 5]:

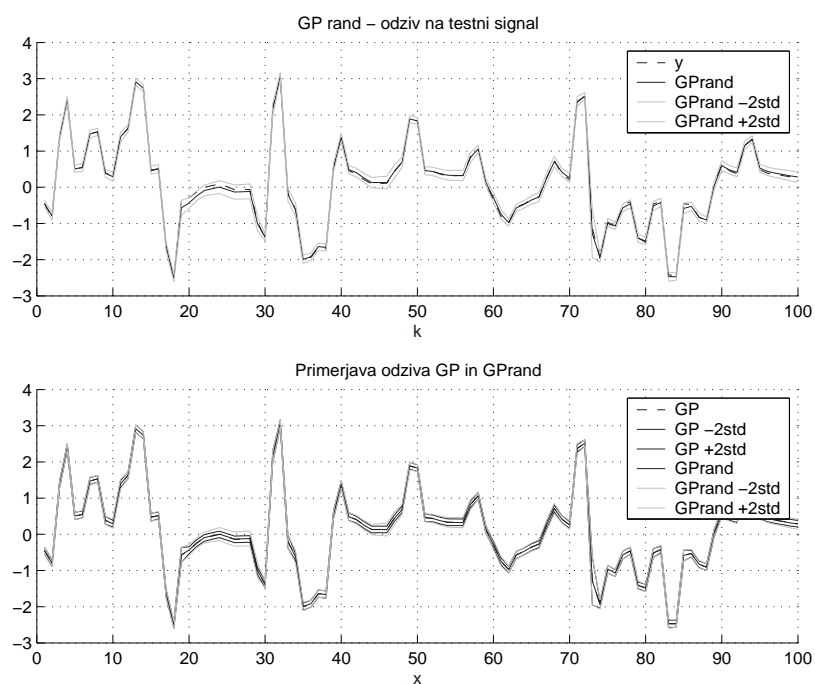
$$\text{Cov}(y(k+n), y(k+n-j)) = \sum_i \beta_i C_{\text{mod}_1}(x^*, \mathbf{x}_i) (c_i - y(k+n-j)) \quad (3.24)$$

kjer je $c_i = (\mathbf{W}^{-1} + \Sigma_{\mathbf{x}}^{-1})^{-1} (\mathbf{W}^{-1} \mathbf{x}_i + \Sigma_{\mathbf{x}}^{-1} \boldsymbol{\nu})$

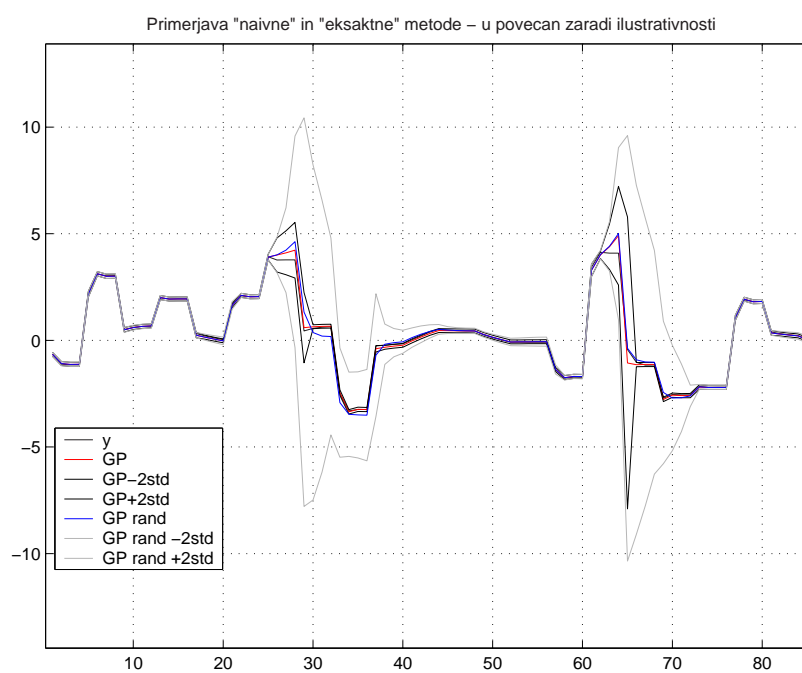
3.5 Primerjava med modeli

3.5.1 Dinamični sistem prvega reda

Na sliki 3.18 je prikazana primerjava simulacije z GPi, kjer v prvem primeru peljemo nazaj na vhod samo srednjo vrednost izhoda ("naivna" metoda, označeno z GP) in v drugem primeru celotno napovedano porazdelitev ("eksaktna", označeno z GPrand). Dokler se vhod nahaja na področju, kjer so se nahajale učne točke, je razlika malo opazna. Na sliki 3.19 smo vhod $u(k)$ povečali, da smo zašli iz področja identifikacije. Tu je razlika med napovedanima negotovostima mnogo bolj očitna.



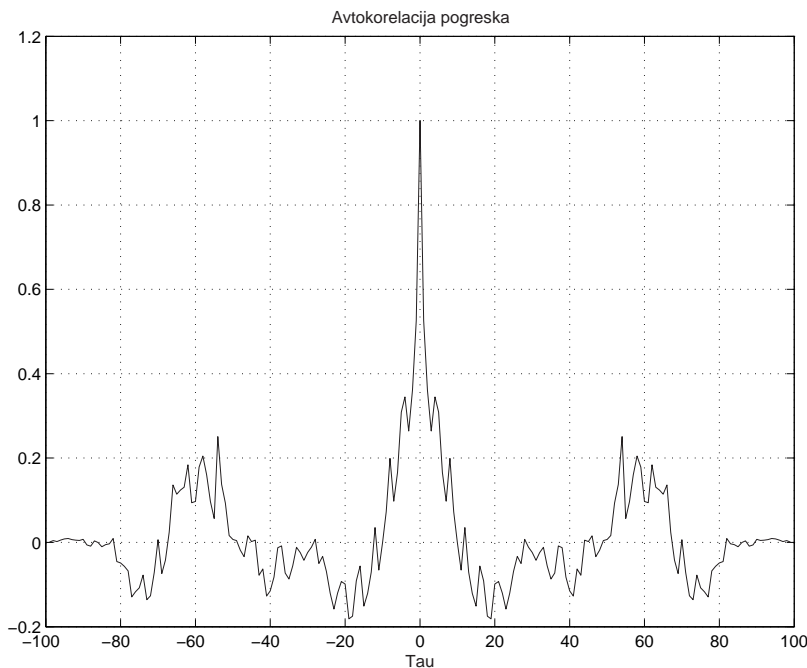
Slika 3.18: Primerjava GPov z in brez uporabe variance na vhodu



Slika 3.19: Primerjava GPov z in brez uporabe variance na vhodu - razširjeno področje

Na slikah 3.20 in 3.21 sta prikazani avtokorelacija pogreška in križna korelacija med vzbujevalnim signalom in pogreškom. Vrednosti cenilk odziva GP modela s propagacijo variance sta $SE = 0.0205$ in $LD = -1.3503$.

Prednost tovrstnega modela je, da kot izhod modela dobimo tudi njegovo negotovost, kar lahko uporabimo na primer za robustno vodenje.



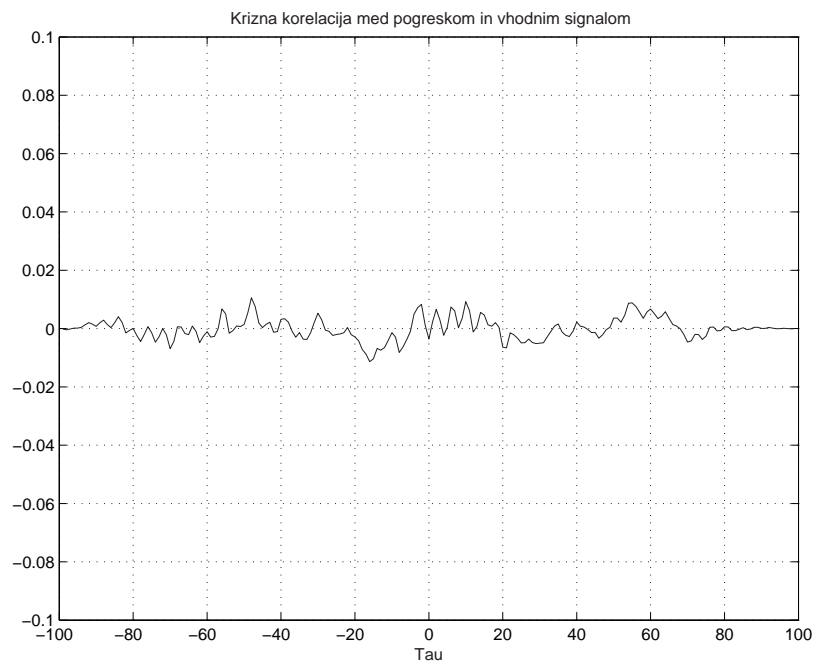
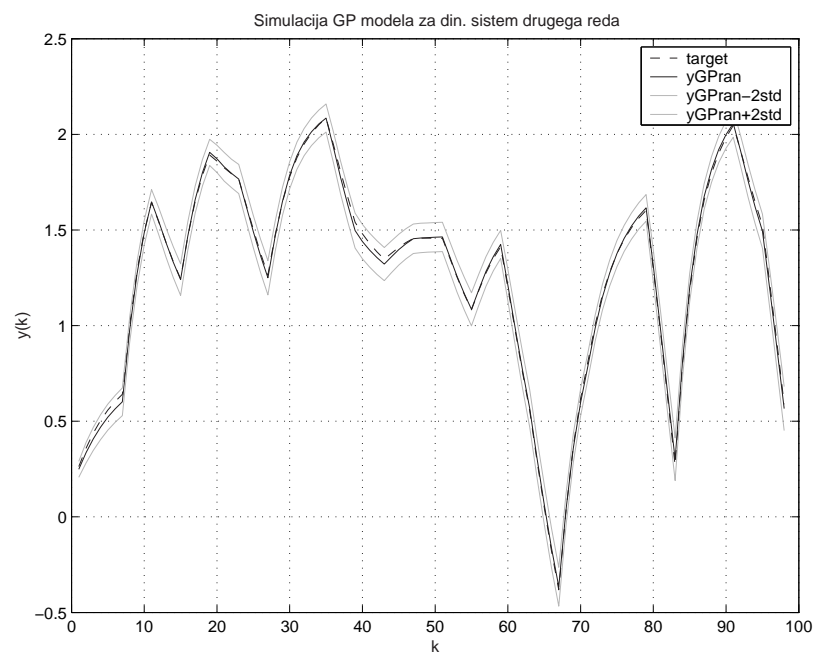
Slika 3.20: Avtokorelacija napake Φ_{ee}

3.5.2 Dinamični sistem drugega reda

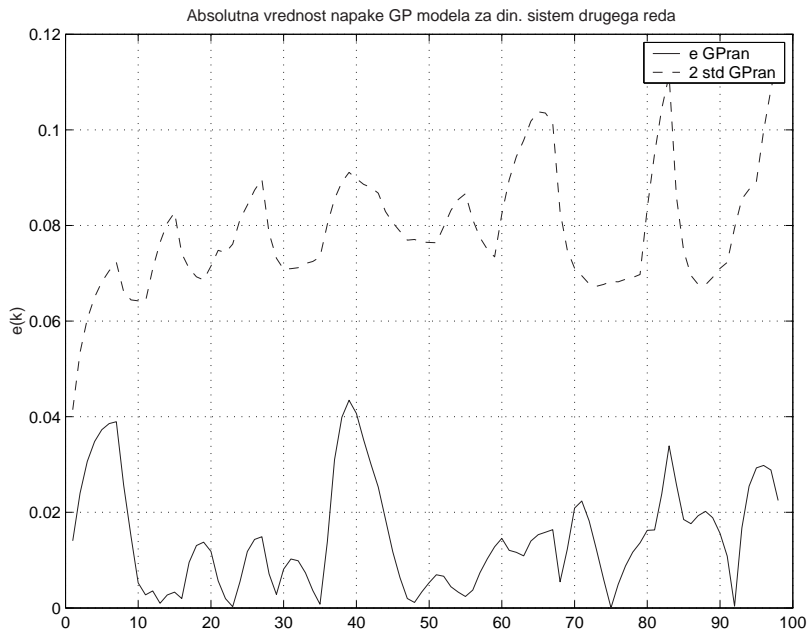
Enako vrednotenje kot v prejšnjem razdelku smo izvedli še za identifikacijo dinamičnega sistema drugega reda.

Rezultati identifikacije nelinearnega dinamičnega sistema (3.3) z "eksaktnim" GP modelom so prikazani na slikah 3.22, 3.23, 3.24 in 3.25. Vrednosti cenilk odziva GP modela s propagacijo variance za dinamični sistem drugega reda sta $SE = 3.3E - 004$ in $LD = -2.22$.

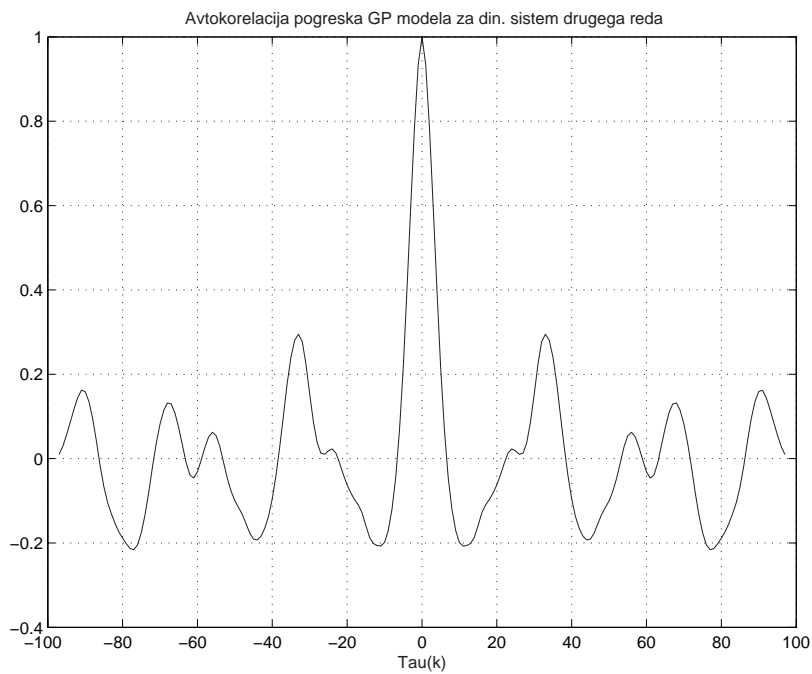
"Eksaktni" GP model izraža manjše zaupanje v predikcijo predvsem na območjih, kjer se spremeni vhodni signal v sistem. V območju, kjer je dovolj učnih podatkov se predikciji modelov in velikosti zaupanja ne razlikujeta veliko.

Slika 3.21: Križna korelacija med vhodom in napako Φ_{ue} 

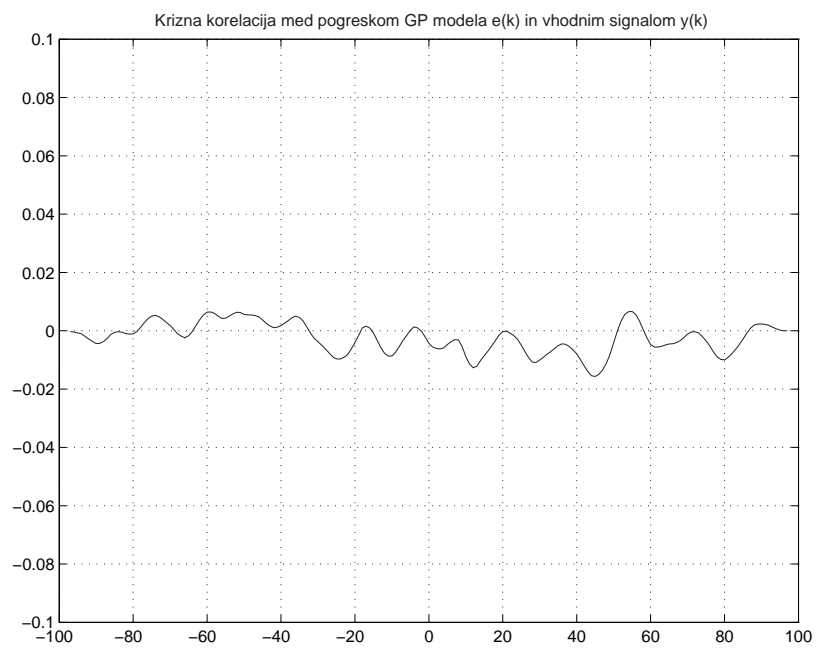
Slika 3.22: Simulacija "eksaktnega" GP modela dinamičnega sistema drugega reda



Slika 3.23: Napaka predikcije eksaktnega GP modela skupaj z zaupanjem



Slika 3.24: Avtokorelacija napake Φ_{ee}

Slika 3.25: Križna korelacija med vhodom in napako Φ_{ue}

4. Vključevanje linearnih lokalnih modelov v GP modele

V prejšnjih razdelkih smo si pogledali, kako lahko z Gaussovimi procesi identificiramo nek nelinearni dinamični sistem. Kljub dobrim lastnostim (metoda poleg napovedane vrednosti vrne tudi zaupanje v to vrednost, dovolj lahka uporaba kljub zapleteni teoriji) ima metoda nekaj pomanjkljivosti. Prva pomanjkljivost je, da za dobro modeliranje nekega hiper-prostora dimenzije D rabimo veliko učnih točk. Število učnih točk pa neposredno vpliva na hitrost metode, saj moramo v vsakem koraku optimizacije parametrov računati inverz kovariančne matrike \mathbf{K} , kjer je število operacij sorazmerno tretji potenci števila učnih točk [2].

Da zagotovimo dobro identifikacijo, morajo biti učni podatki porazdeljeni po vsem prostoru, kjer to želimo. Običajno imamo na voljo veliko podatkov v okolici ravnotežnih točk in relativno malo podatkov daleč od ravnotežnih točk. Več o tem problemu najdemo v [16].

Identifikacija z vpeljavo predznanja v obliki linearnih lokalnih modelov (LM) v GPe (krajše LM GP model) poizkusi obdržati dobre in odpraviti nekatere slabe lastnosti identifikacije z GPi.

Ideja vpeljave linearnih lokalnih modelov (v nadaljevanju lokalnih modelov) je, da veliko število podatkov, ki so dovolj blizu skupaj (ponavadi v okolici ravnotežnih točk) predstavimo z determinističnim modelom linearnega dinamičnega sistema. S tem lahko uspešno zmanjšamo število učnih točk, ki naj opisujejo neko področje znotraj področja, ki ga hočemo identificirati.

Postavi se vprašanje, če ne bi bilo uporabno kar celotno delovno področje predstaviti kot mrežo lokalnih modelov. Ta bi verjetno odlično predstavljala sistem po vsem identificiranem področju, vendar se izven bližnjega območja ravnotežne krivulje pojavi problem pridobivanja lokalnih modelov, saj imamo za to običajno

premalo podatkov.

Za zgraditev linearnega modela so potrebne tako izmerjene funkcijske vrednosti (ang. functional observations, v nadaljevanju funkcijske vrednosti) pri določenem vhodu $y = f(\mathbf{x})$, kot tudi (parcialni) odvodi funkcije (ang. derivative observations, v nadaljevanju izmerjeni odvodi) po komponentah vhodnega vektorja $\frac{\partial f}{\partial x_k}$. Izmerjene odvode lahko uporabimo, saj je odvajanje linearna operacija in odvod GPa ostane GP [22].

Ker imamo sedaj nov tip podatkov, se mora za te tipe spremeniti tudi kovariančna funkcija [8]. Vektor zelenih izhodov (ang. target vector) \mathbf{y} , ki je bil prej samo vektor meritev izhoda, sedaj vsebuje tudi vrednosti parcialnih odvodov v določenih točkah. Vrednostim odvoda v zelenem izhodu na vhodu ustrezajo vrednosti regresorjev za odgovarjajoč izmerjen odvod.

Več podrobnosti o vključevanju LM v GP lahko najdemo v [8, 22]. Opozoriti velja na še en način vključevanja predznanja o LM, ki ga pa ne bomo vrednotili. Pri tem pristopu več lokalnih modelov, ki so realizirani z GPi (*Gaussovi lokalni modeli*) lahko združimo v globalni model [7].

Vzemimo ponovno sistem, opisan z enačbo (3.1):

$$y(k) = f(y(k-1), \dots, y(k-n), u(k-1), \dots, u(k-n)) \quad (4.1)$$

Ena možnost (vsekakor ne edina!) je, da si učne podatke – matriko vhodov dimenzije $N \times D$, $D = 2L$, kjer je L red sistema – in vektor zelenih izhodov dolžine N , izberemo kot:

$$\mathbf{X} = \begin{bmatrix} \mathbf{Y}_{oeq} & \mathbf{U}_{oeq} \\ \mathbf{Y}_{eq} & \mathbf{U}_{eq} \\ \mathbf{Y}_{eq} & \mathbf{U}_{eq} \\ \vdots & \vdots \\ \mathbf{Y}_{eq} & \mathbf{U}_{eq} \\ \vdots & \vdots \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} \mathbf{Y}_{oeq1} \\ \mathbf{Y}_{eq1} \\ [\frac{\partial f}{\partial y(k)}] \\ \vdots \\ [\frac{\partial f}{\partial u(k)}] \\ \vdots \end{bmatrix} \quad (4.2)$$

kjer so:

\mathbf{Y}_{oeq1} vektor dolžine n_{oeq} zelenih odzivov izven ravnotežne krivulje;

\mathbf{Y}_{oeq} vhodna matrika velikosti $n_{oeq} \times L$ preteklih izhodov izven ravnotežne krivulje;

\mathbf{U}_{oeq} vhodna matrika velikosti $n_{oeq} \times L$ vzbujevalnih signalov;

\mathbf{Y}_{eq1} vektor dolžine n_{eq} odzivov na ravnotežni krivulji;

\mathbf{Y}_{eq} matrika velikosti $n_{eq} \times L$ odzivov na ravnotežni krivulji;

\mathbf{U}_{eq} matrika velikosti $n_{eq} \times L$ vzbujevalnih vhodov na ravnotežni krivulji;

$[\frac{\partial f}{\partial y}]$ vektor dolžine n_D izmerjenih odvodov po zakasnenih vrednostih izhodnega signala (vektor koeficientov linearnega modela v različnih točkah);

$[\frac{\partial f}{\partial u}]$ vektor dolžine n_D izmerjenih odvodov po zakasnenih vrednostih vzbujevalnega signala (vektor koeficientov linearnega modela v različnih točkah).

Skupaj imamo D vektorjev odvodov ($[\frac{\partial f}{\partial y_1}]$, ... , $[\frac{\partial f}{\partial y_L}]$, $[\frac{\partial f}{\partial u_1}]$, ... , $[\frac{\partial f}{\partial u_L}]$), po enega za posamezen regresor. Velikost matrike učnih vhodov je: $N = (n + D \cdot n_D) \times D$, kjer je $n = n_{oeq} + n_{eq}$, $n_D = n_{eq}$, velikost vektorja učnih izhodov pa $(n + D \cdot n_D) \times 1$.

4.1 Kovariančna funkcija

Ker se je spremenila narava podatkov, se mora spremeniti tudi kovariančna funkcija [8]. Kot vhod vanjo še vedno nastopajo vhodni vektorji \mathbf{x}_i , sestavljeni iz vrednosti regresorjev $[y(k-1), \dots, y(k-n), u(k-1), \dots, u(k-n)]$ dimenzije D . Kovariančna funkcija med dvema funkcijskima vrednostima ostane enaka:

$$C(\mathbf{x}_i, \mathbf{x}_j) = v \exp \left[-\frac{1}{2} \sum_{d=1}^D w_d (x_i^d - x_j^d)^2 \right] \quad (4.3)$$

Kovariančna funkcija med funkcijsko vrednostjo in izmerjenim odvodom:

$$C\left(\frac{\partial \mathbf{x}_i}{\partial x_i}, \mathbf{x}_j\right) = -v w_d (x_i^d - x_j^d) \exp \left[-\frac{1}{2} \sum_{d=1}^D w_d (x_i^d - x_j^d)^2 \right] \quad (4.4)$$

kjer je d indeks odvajanega regresorja vhoda \mathbf{x}_i .

Kovariančna funkcija med dvema odvodoma:

$$C\left(\frac{\partial \mathbf{x}_i}{\partial x_i}, \frac{\partial \mathbf{x}_j}{\partial x_j}\right) = v w_e (\delta_{e,d} - w_d (x_i^e - x_j^e) (x_i^d - x_j^d)) \exp \left[-\frac{1}{2} \sum_{d=1}^D w_d (x_i^d - x_j^d)^2 \right] \quad (4.5)$$

kjer je $\delta_{e,d}$ Kronecker-jev operator med e -to komponento (e -ti regresor) vektorja \mathbf{x}_i in d -to komponento (d -ti regresor) vektorja \mathbf{x}_j .

Za učne podatke, ki so zloženi kot v zgornjem primeru (glej enačbo 4.2), je kovariančna matrika sestavljena:

$$\mathbf{K} = \begin{bmatrix} \left[C(\mathbf{x}_i, \mathbf{x}_j) \right] & \left[C(\mathbf{x}_i, \frac{\partial \mathbf{x}_j}{\partial x_j}) \right]_{d=1} & \cdots & \left[C(\mathbf{x}_i, \frac{\partial \mathbf{x}_j}{\partial x_j}) \right]_{d=D} \\ \left[C(\frac{\partial \mathbf{x}_i}{\partial x_i}, \mathbf{x}_j) \right]_{d=1} & \left[C(\frac{\partial \mathbf{x}_i}{\partial x_i}, \frac{\partial \mathbf{x}_j}{\partial x_j}) \right]_{e=1,d=1} & \cdots & \left[C(\frac{\partial \mathbf{x}_i}{\partial x_i}, \frac{\partial \mathbf{x}_j}{\partial x_j}) \right]_{e=1,d=D} \\ \vdots & \vdots & \vdots & \vdots \\ \left[C(\frac{\partial \mathbf{x}_i}{\partial x_i}, \mathbf{x}_j) \right]_{d=D} & \left[C(\frac{\partial \mathbf{x}_i}{\partial x_i}, \frac{\partial \mathbf{x}_j}{\partial x_j}) \right]_{e=D,d=1} & \cdots & \left[C(\frac{\partial \mathbf{x}_i}{\partial x_i}, \frac{\partial \mathbf{x}_j}{\partial x_j}) \right]_{e=D,d=D} \end{bmatrix} \quad (4.6)$$

Meritve lahko vključujejo informacijo o šumu. Če je informacija vključena, njeno vrednost dodamo ustreznim diagonalnim elementom v kovariačni matriki. Če te informacije nimamo, učimo parameter v_0 kot je opisano v poglavju 2.4.

Ko merimo odvode, dobimo pri uporabi standardnih metod za identifikacijo hkrati z izmerjenimi odvodi tudi informacijo o šumu, ki jo vključimo v kovariančno matriko [22].

Izhod iz modela – srednja vrednost in varianca normalne porazdelitve – se izračuna kot za navadni model GP (enačbi 2.8 in 2.9).

Več informacij o lokalnih modelih, vključenih v GP modele, se nahaja v [8, 22].

4.2 Predikcija pri porazdelitvi kot vhodu v LM GP model

Ideja je ista kot pri razširitvi normalnega GP modela. Namesto srednje vrednosti naj se iz izhoda na vhod prenese celotna (predpostavljamo normalna) porazdelitev. Tako izveden model izraža stopnjo zaupanja v napovedano vrednost (analogno kot pri navadnih modelih GP).

Izhajamo iz enačb (3.6) in (3.7). Po [5] oziroma razdelku 3.4 se "nova" napovedana srednja vrednost $m(\cdot)$ in njena varianca $v(\cdot)$ računata kot:

$$m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = E_{\mathbf{x}}[\mu(\mathbf{x})] \quad (4.7)$$

$$v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = E_{\mathbf{x}}[\sigma^2(\mathbf{x})] + E_{\mathbf{x}}[\mu(\mathbf{x})^2] - E_{\mathbf{x}}[\mu(\mathbf{x})]^2 \quad (4.8)$$

4.3 Napovedana srednja vrednost za nov $\mathbf{x} \sim \mathcal{N}_{\mathbf{x}}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$

Vsak regresor je sedaj predstavljen kot porazdelitev. Vektor $\boldsymbol{\nu}$ predstavlja srednje vrednosti regresorjev, matrika $\boldsymbol{\Sigma}_x$ pa vrednosti njihovih križnih korelacij.

Izpeljavo za ta in naslednji razdelek najdemo v [8], tu so zaradi obsega podani le končni rezultati. Osnovna ideja je računanje kovariančne funkcije posebej med izmerjenimi funkcijskimi vrednostmi in testnim vhomom ter izmerjenimi odvodi in testnim vhomom.

Za napoved srednje vrednosti moramo izračunati $m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = E_{\mathbf{x}}[\mu(\mathbf{x})]$, to je

$$\boxed{m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = E_{\mathbf{x}}[\mu(\mathbf{x})] = E_{\mathbf{x}}[\mu_1(\mathbf{x})] + E_{\mathbf{x}}[\mu_d(\mathbf{x})]} \quad (4.9)$$

Rešitev za del, ki predstavlja prispevek funkcijskih vrednosti je že znana iz enačbe (3.18):

$$E_{\mathbf{x}}[\mu_1(\mathbf{x})] = m_1(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = \sum_i \beta_i C_{mod_1}(\boldsymbol{\nu}, \mathbf{x}_i) \quad (4.10)$$

kjer so β_i elementi vektorja $\boldsymbol{\beta} = \mathbf{K}^{-1}\mathbf{y}$ in C_{mod_1} opisuje (4.11).

$$C_{mod_1}(\mathbf{x}_i, \boldsymbol{\nu}) = v_0 |\mathbf{I} + \mathbf{W}^{-1}\boldsymbol{\Sigma}_x|^{-1/2} \exp \left[-\frac{1}{2} (\mathbf{x}_i - \boldsymbol{\nu})^T (\mathbf{W} + \boldsymbol{\Sigma}_x)^{-1} (\mathbf{x}_i - \boldsymbol{\nu}) \right] \quad (4.11)$$

Vrednost $E_{\mathbf{x}}[\mu_d(\mathbf{x})]$, ki predstavlja prispevek izmerjenih odvodov pa se izračuna:

$$E_{\mathbf{x}}[\mu_d(\mathbf{x})] = - \sum_{d=1}^D w_d \sum_i \beta_i (x_i^d - c_i^d) C_{mod_1}(\mathbf{x}_i, \boldsymbol{\nu}) \quad (4.12)$$

kjer je c_i^d d -ta komponenta vektorja $(\mathbf{W}^{-1} + \boldsymbol{\Sigma}_x^{-1})^{-1} (\mathbf{W}^{-1}\mathbf{x}_i + \boldsymbol{\Sigma}_x^{-1}\boldsymbol{\nu})$, x_i^d pa d -ta komponenta učnega vektorja \mathbf{x}_i .

4.4 Napovedana varianca za nov $\mathbf{x} \sim \mathcal{N}_{\mathbf{x}}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$

Napovedana varianca se izpelje iz izraza (4.8).

Njen izraz vsebuje elemente varianc funkcijskih vrednosti (indeks 1), varianc izmerjenih odvodov (indeks d) in varianc mešanih izrazov (indeks $1d$):

$$\boxed{v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = v + E_{\mathbf{x}}[\sigma_1^2(\mathbf{x})] + 2E_{\mathbf{x}}[\sigma_{1d}^2(\mathbf{x})] + E_{\mathbf{x}}[\sigma_d^2(\mathbf{x})] + E_{\mathbf{x}}[\mu_1(\mathbf{x})^2] + 2E_{\mathbf{x}}[\mu_{1d}(\mathbf{x})^2] + E_{\mathbf{x}}[\mu_d(\mathbf{x})^2] - m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)^2} \quad (4.13)$$

ki se izračunajo kot:

$$E_{\mathbf{x}}[\sigma_1^2(\mathbf{x})] = - \sum_{i,j} K_{ij}^{-1} C_{mod_2}(\mathbf{x}_i, \mathbf{x}_j) C_{mod_3}(\boldsymbol{\nu}, \mathbf{x}_b) \quad (4.14)$$

$$E_{\mathbf{x}}[\sigma_{1d}^2(\mathbf{x})] = \sum_{d=1}^D w_d \sum_{i,j} K_{ij}^{-1} [x_i^d - c_{ij}^d] C_{mod_2}(\mathbf{x}_i, \mathbf{x}_j) C_{mod_3}(\boldsymbol{\nu}, \mathbf{x}_b) \quad (4.15)$$

kjer $\mathbf{x}_b = \frac{\mathbf{x}_i + \mathbf{x}_j}{2}$.

$$\begin{aligned} E_{\mathbf{x}}[\sigma_d^2(\mathbf{x})] = & \\ & - \sum_{e,d=1}^D w_e w_d \sum_{i,j} K_{ij}^{-1} (x_i^e x_j^d - (x_i^e c_{ij}^d + x_j^d c_{ij}^e) + C_{ed} + c_{ij}^e c_{ij}^d) C_{mod_2}(\mathbf{x}_i, \mathbf{x}_j) C_{mod_3}(\boldsymbol{\nu}, \mathbf{x}_b) \end{aligned} \quad (4.16)$$

kjer je C_{ed} (e, d)-ti element matrike $\mathbf{C} = \left(\left(\frac{\mathbf{W}}{2} \right)^{-1} + \boldsymbol{\Sigma}_x^{-1} \right)^{-1}$ in c_{ij}^d d -ti element vektorja $\mathbf{C} \left(\left(\frac{\mathbf{W}}{2} \right)^{-1} \frac{\mathbf{x}_i + \mathbf{x}_j}{2} + \boldsymbol{\Sigma}_x^{-1} \boldsymbol{\nu} \right)$.

$$E_{\mathbf{x}}[\mu_1(\mathbf{x})^2] = \sum_{i,j} \beta_i \beta_j C_{mod_2}(\mathbf{x}_i, \mathbf{x}_j) C_{mod_3}(\mathbf{x}_i, \mathbf{x}_j) \quad (4.17)$$

$$E_{\mathbf{x}}[\mu_{1d}(\mathbf{x})^2] = - \sum_{d=1}^D w_d \sum_{i,j} \beta_i \beta_j [x_i^d - c_{ij}^d] C_{mod_2}(\mathbf{x}_i, \mathbf{x}_j) C_{mod_3}(\boldsymbol{\nu}, \mathbf{x}_b) \quad (4.18)$$

$$E_{\mathbf{x}}[\mu_d(\mathbf{x})^2] = \sum_{e,d=1}^D w_e w_d \sum_{i,j} \beta_i \beta_j (x_i^e x_j^d - (x_i^e c_{ij}^d + x_j^d c_{ij}^e) + C_{ed} + c_{ij}^e c_{ij}^d) C_{\text{mod}_2}(\mathbf{x}_i, \mathbf{x}_j) C_{\text{mod}_3}(\boldsymbol{\nu}, \mathbf{x}_b) \quad (4.19)$$

Izraza za C_{mod_2} in C_{mod_3} sta podana že v enačbah (3.21) in (3.22).

4.5 Simulacija

Postopek simulacije modela je enak kot za GP s porazdelitvijo kot vhodom. Opisano je v razdelku 3.4, le da sedaj za računanje križne kovariance med izhodi, ki jih pripeljemo nazaj na vhod ne moremo uporabiti izraza (3.24).

Za sistem prvega reda ($L = 1$) se postopek računanja matrike križnih kovarianc iz enačbe (3.23) poenostavi in matrika križnih kovarianc degenerira v:

$$x_{t+k} \sim \mathcal{N}(m(x_{t+k-1}), v(x_{t+k-1}) + v_0)$$

kjer križnih kovarianc sploh ni potrebno računati.¹

4.6 Programska oprema za identifikacijo z vključenimi lokalnimi modeli

Programska oprema, uporabljena za identifikacijo z vključenimi lokalnimi modeli, je bila pisana za programsko okolje *Matlab*. Uporabljali smo metode različnih avtorjev in med sabo niso bile povsem kompatibilne. Rutine je bilo treba preizkusiti in dopisati manjkajoče dele. Med vrednotenjem se je pokazalo, da so imele nekatere programske rešitve pomanjkljivosti, ki jih je bilo potrebno odpraviti. Kljub vložnemu trudu pa ostaja konsistentna in učinkovita programska oprema za identifikacijo dinamičnih sistemov z GPi naloga za prihodnost.

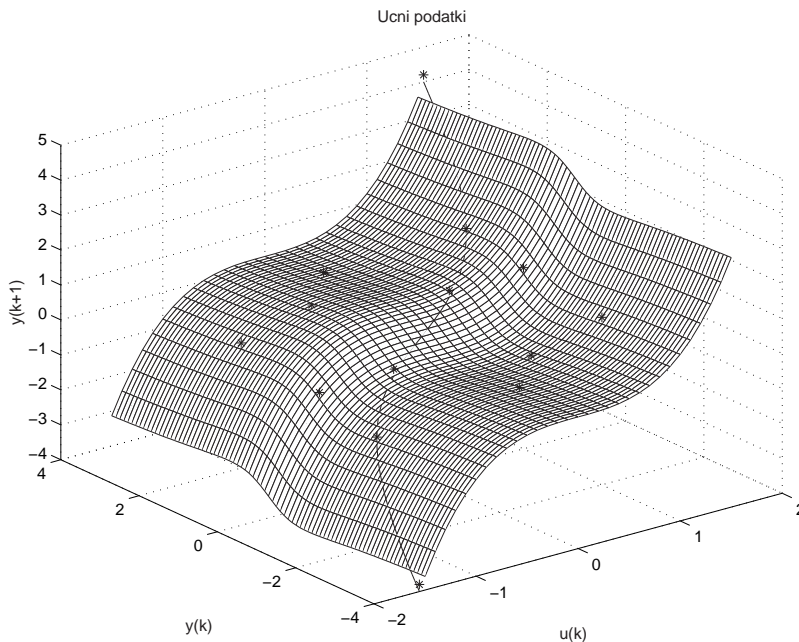
¹Zakaj omenjamo poenostavitev za prvi red bo razvidno pri vrednotenju sistema drugega reda.

4.7 Primera vključevanja lokalnih modelov v GP model

Za prikaz vključevanja lokalnih modelov v GP model smo si izbrali sistema, ki sta že bila identificirana z GP modelom v razdelku 3.2

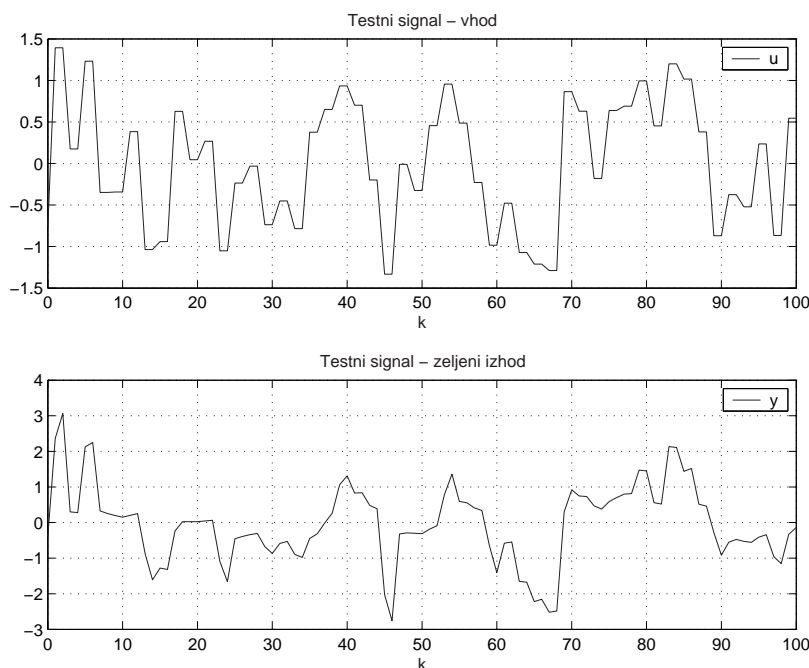
4.7.1 Dinamični sistem prvega reda

Sistem opisuje enačba (3.2). Učili smo ga z dvema učnima množicama podatkov. Prvo učno množico je sestavljalo šest lokalnih modelov (za vsakega od dveh regresorjev po šest izmerjenih odvodov – $n = 6, n_D = 6$) in šest funkcijskih vrednosti ravnotežnih točk, v katerih smo poiskali lokalne modele. V drugi učni množici smo podatkom iz prve učne množice dodali še šest funkcijskih točk izven ravnotežne krivulje ($n = 12, n_D = 6$). Na sliki 4.1 so prikazani učni podatki za oba primera. Podatki za testiranje so bili dobljeni enako kot za identifikacijo GP modela brez vključenih lokalnih modelov (3.3.1) in so prikazani na sliki 4.2.



Slika 4.1: Učni podatki

Funkcijske vrednosti izven ravnotežne krivulje so vsebovale beli šum z varianco $\sigma^2 = 4 \cdot 10^{-4}$. Za določitev izmerjenih odvodov (koeficientov linearnega modela) je bil uporabljen algoritem pomožnih spremenljivk iz programskega paketa Matlab



Slika 4.2: Signal za vrednotenje LM GP modela za din. sistem prvega reda

(IV_4), ki da poleg vrednosti odvodov še njihovo kovariančno matriko. Za merjenje odvodov v ravnotežnih točkah je bil uporabljen psevdo-naključni binarni signal (PNBS) z amplitudo 0.04, izhodnemu signalu je bil prištet beli šum z največjo vrednostjo $\pm 2.5\%$ amplitude PNBS-a. Parametri identificiranih lokalnih modelov so podani v tabeli 4.1. Variance lokalnih modelov so bile velikostnega reda 10^{-4} in manjše. V delovni točki (U_{eq}, Y_{eq}) opisuje delovanje lokalnega modela enačba:

$$y(k) = -ay(k-1) + bu(k-1) \quad (4.21)$$

Na slikah 4.3, 4.4 in 4.5 si lahko ogledamo odziv, napako in varianco enokoračne predikcije navadne metode z vključenimi LM za prvo učno množico (samo lokalni modeli brez funkcijskih točk izven ravnotežne krivulje), na slikah 4.6, 4.7 in 4.8 pa odziv, napako in varianco za drugo učno množico (lokalni modeli in funkcijske točke). Opazimo lahko, da sta odziv in varianca v okolici ravnotežne krivulje dobra za oba identificirana modela.

Modela se po pričakovanju bolj razlikujeta izven ravnotežne krivulje. Dodatne učne točke, s katerimi smo učili drugi model, močno približajo odziv modela odzivu sistema (za primerjavo odzivov primerjamo sliki 4.3 in 4.6 ali sliki 4.4 in

4.7).

Odziv prvega modela tudi izven središča lokalnega modela ohranja njegov trend, medtem ko se odziv drugega modela lepo prilagodi, ko pride zadosti blizu učni točki. Če vemo približno kakšne vrednosti signalov lahko pričakujemo pri vzbujanju določenega sistema, lahko model s pravo kombinacijo točk izven ravnotežne krivulje identificiramo boljše.

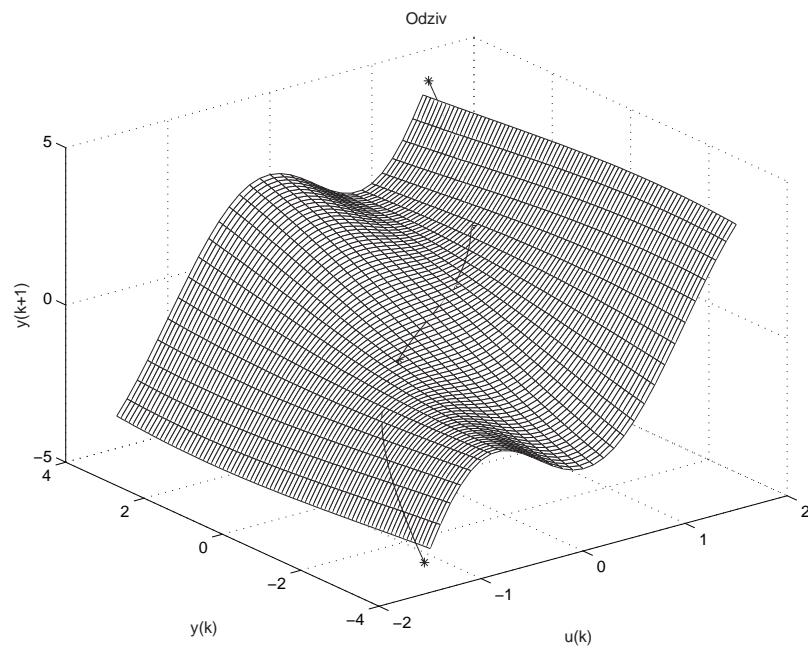
Ko primerjamo varianci obeh modelov (sliki 4.5 in 4.8) opazimo, da predlagana varianca drugega modela na robovih identifikacijskega področja narašča bolj strmo in je v območju učnih točk manjša od variance prvega modela. Varianca prvega modela je večja, saj ima model pri učenju na voljo manj informacije kot drugi.

Na sliki 4.9 je prikazana simulacija prvega modela. Model pri velikih spremembah vhodnega signala ne zmore več dobro napovedovati izhoda, saj so napovedi vhodov, ki so bolj oddaljeni od ravnotežne krivulje, premalo natančni. Če pa bi imeli počasno se spreminjajoč signal, bi to pomenilo majhno odstopanje od ravnotežne krivulje in dovolj natančno napovedovanje tudi samo z informacijo lokalnih modelov.

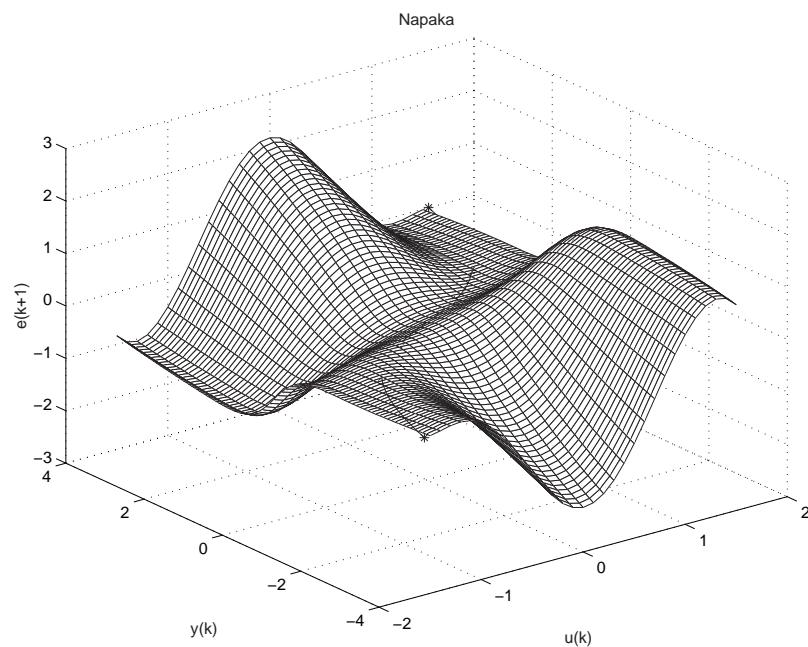
Na sliki 4.10 je prikazana simulacija drugega modela z "naivno" metodo, na sliki 4.11 pa sta skupaj prikazana rezultata simulacije z "naivno" in "eksaktno" metodo. Modela sta zadovoljivo identificirala neznan sistem, zaradi dobre izbire učnega področja se tudi zaupanji v vrednosti izhoda malo razlikujeta. Na sliki 4.12 sta prikazani napaki simulacije obeh metod. Slika 4.13 prikazuje

Tabela 4.1: Parametri uporabljenih lokalnih modelov za dinamični sistem prvega reda

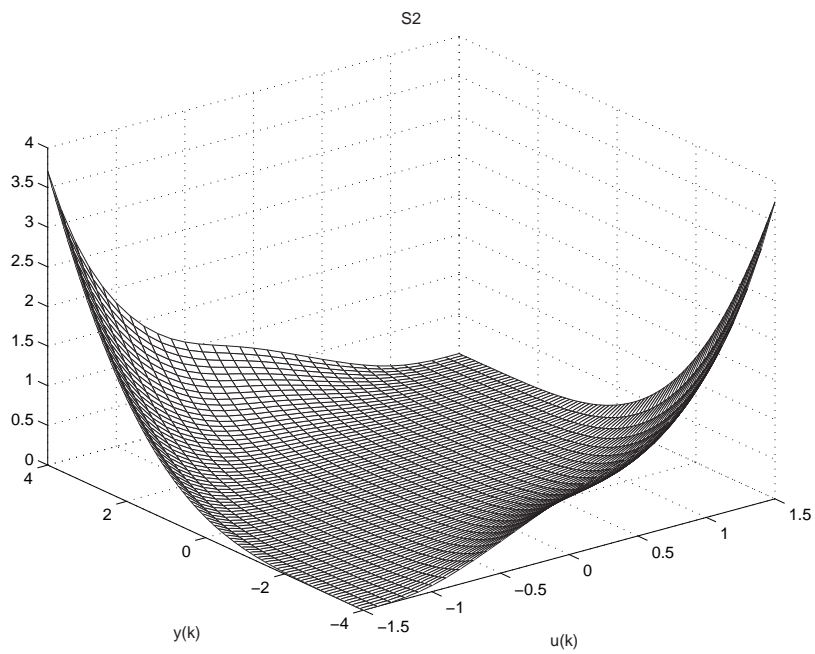
U_{eq}	Y_{eq}	$-a$	b
-1.5556	-4.0000	-0.0518	7.2583
-1.0127	-1.5000	-0.1175	3.0786
-0.4642	-0.5000	0.4736	0.6473
0.4642	0.5000	0.4773	0.6480
0.9751	1.4000	-0.1094	2.8548
1.5556	4.0000	-0.0517	7.2645



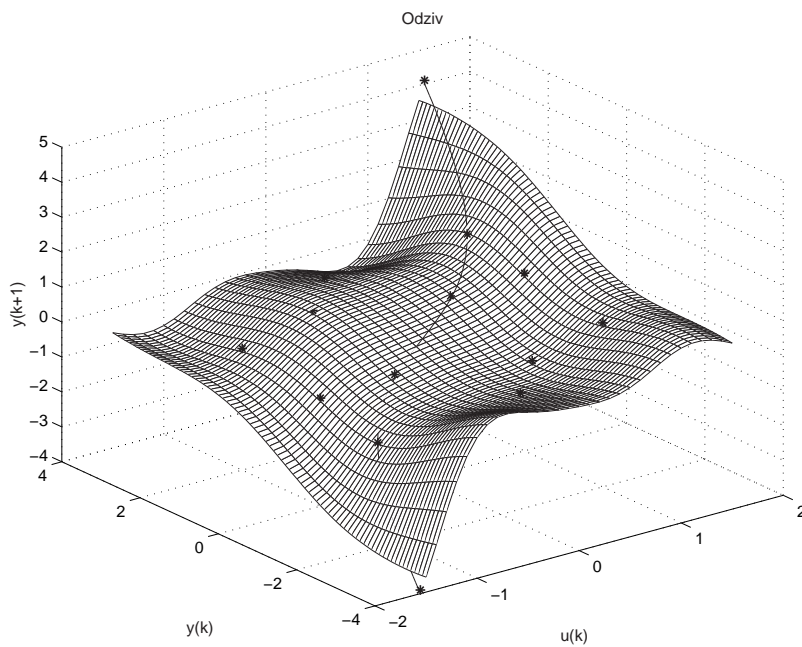
Slika 4.3: Enokoračna predikcija (samo lokalni modeli)



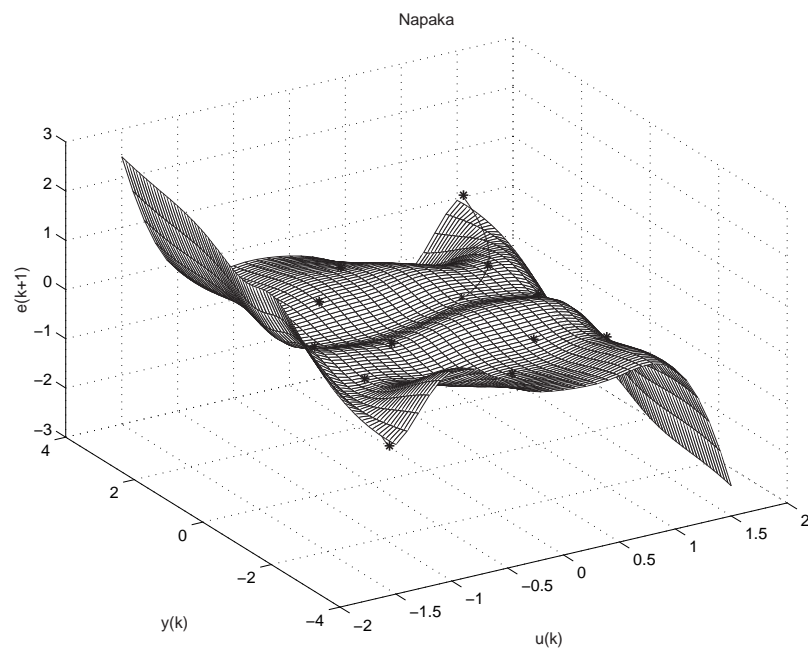
Slika 4.4: Napaka enokoračne predikcije (samo lokalni modeli)



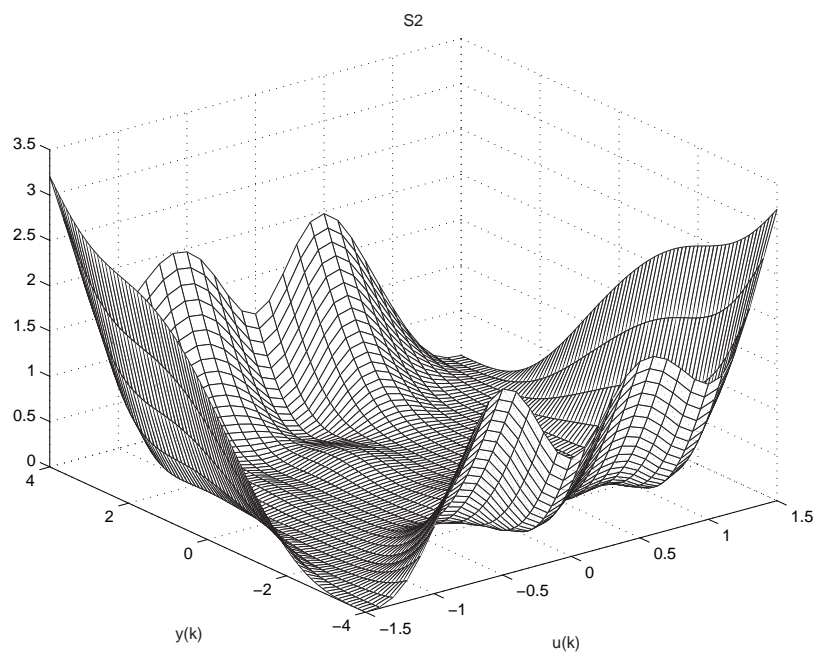
Slika 4.5: Varianca (samo lokalni modeli)



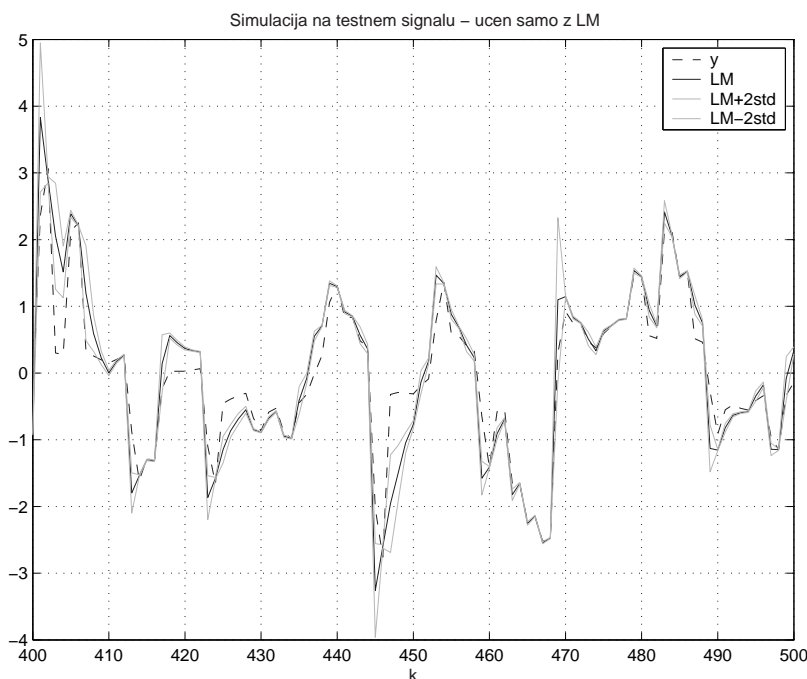
Slika 4.6: Enokoračna predikcija



Slika 4.7: Napaka enokoračne predikcije



Slika 4.8: Varianca



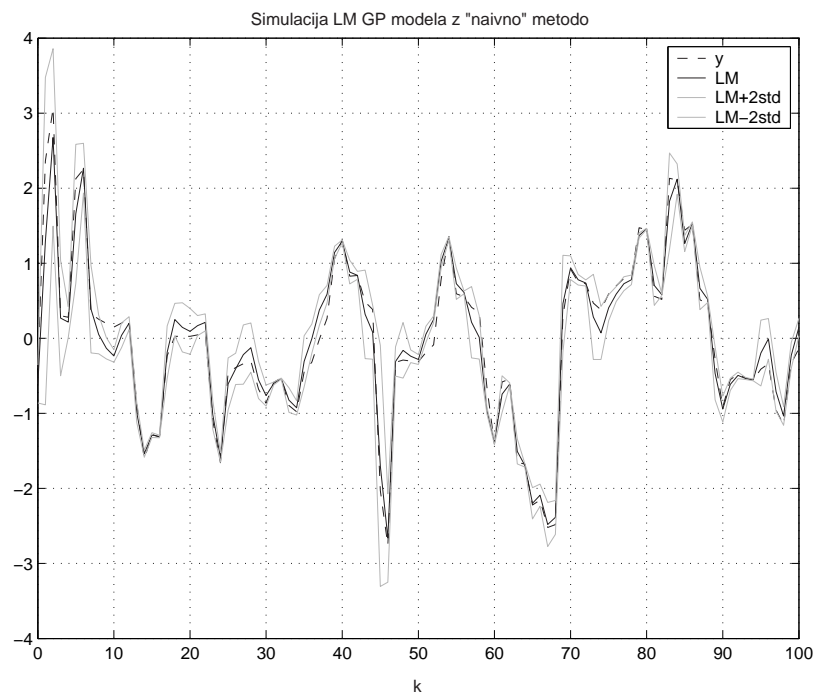
Slika 4.9: Simulacija (samo lokalni modeli)

povečanje variance, ko z modelom preidemo na neznano področje. Velikost vhoda u je bila namerno povečana, da se bolje izrazi razlika med napovedanim zaupanjem "naivne" in "eksaktne" metode.

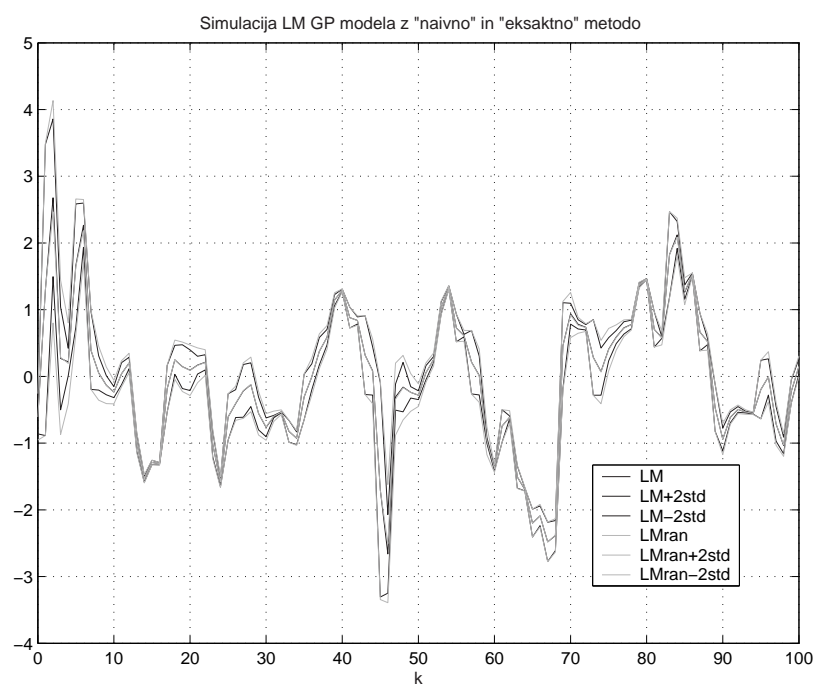
Statistični oceni obeh modelov – avtokovarianca napake $\Phi_{ee}(\tau)$ na sliki 4.14 in križna kovarianca med vhodom $u(k)$ in napako $e(k+1)$ $\Phi_{ue}(\tau)$ na sliki 4.15 – sta se pokazali za zadovoljivi tako za model, katerega vhod so bile porazdelitve kot za model, katerega vhod so bile samo vrednosti (oceni sta prikazani samo za "naivni" model, saj se praktično ne razlikujeta).

Vrednosti cenilk odziva LM GP modela brez propagacije variance sta $SE = 0.0406$, $LD = 3.5266$, s propagacijo pa $SE = 0.0435$, $LD = 0.0331$.

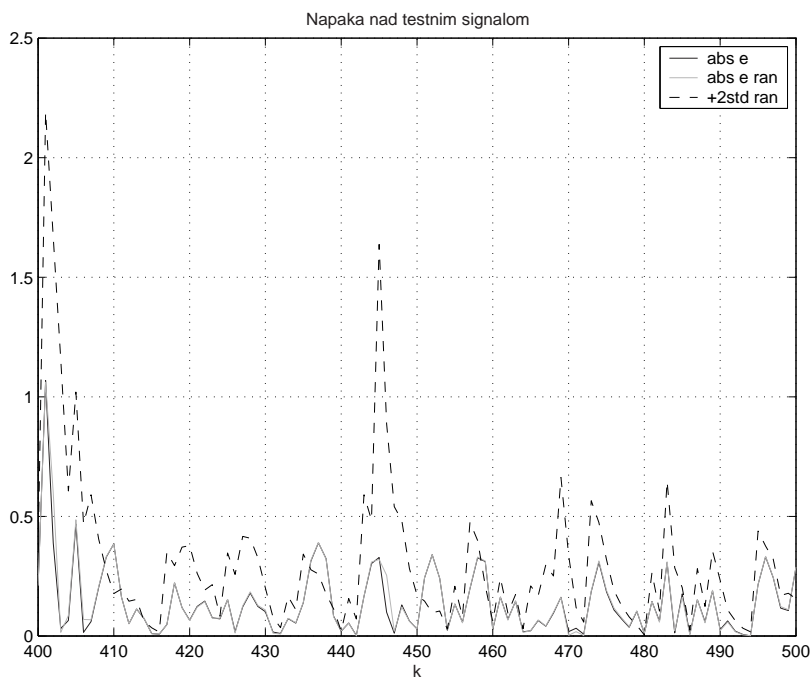
LM GP model je na učnem področju lepo povzel glavne značilnosti sistema kljub velikem zmanjšanju števila učnih podatkov glede na navadni GP model. Primerjava dveh modelov, kjer je bil prvi učen samo z lokalnimi modeli, drugi pa še s funkcijskimi točkami, oddaljenimi od ravnotežne krivulje je pokazala vpliv teh točk na odziv stran od ravnotežne krivulje. Z relativno malo modeli in funkcijskimi točkami smo dobro ujeli glavne značilnosti sistema na obravnavanem področju.



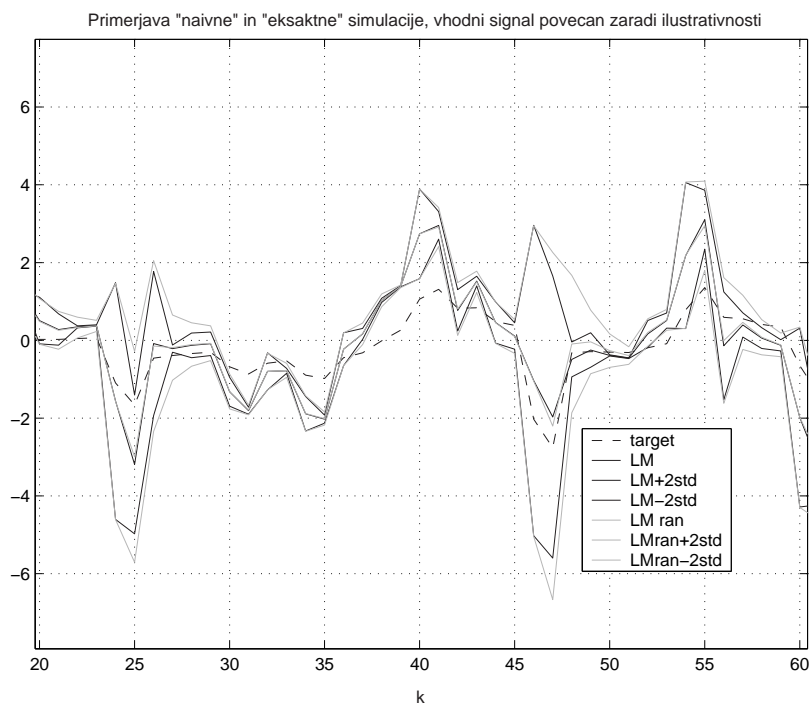
Slika 4.10: Simulacija z "naivno" metodo



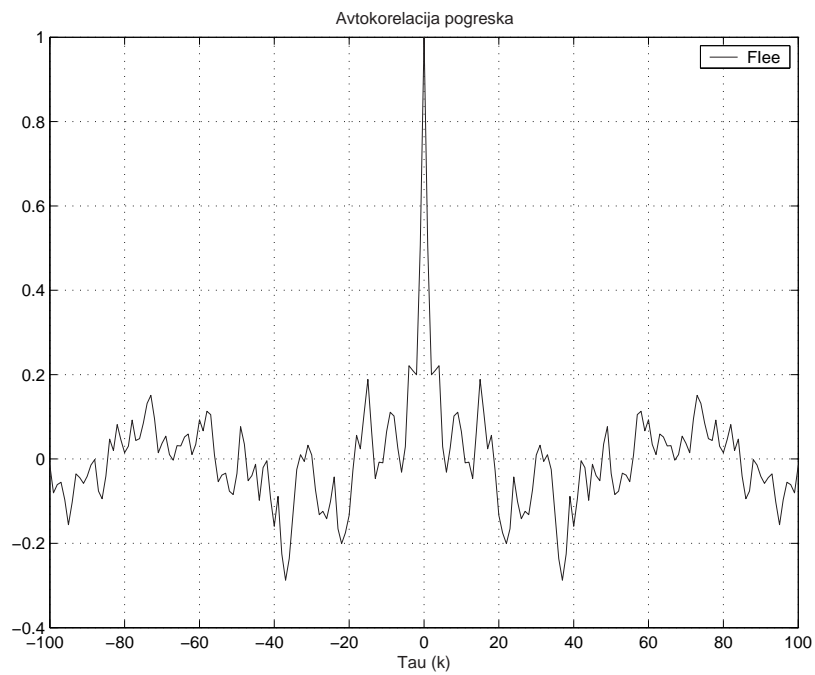
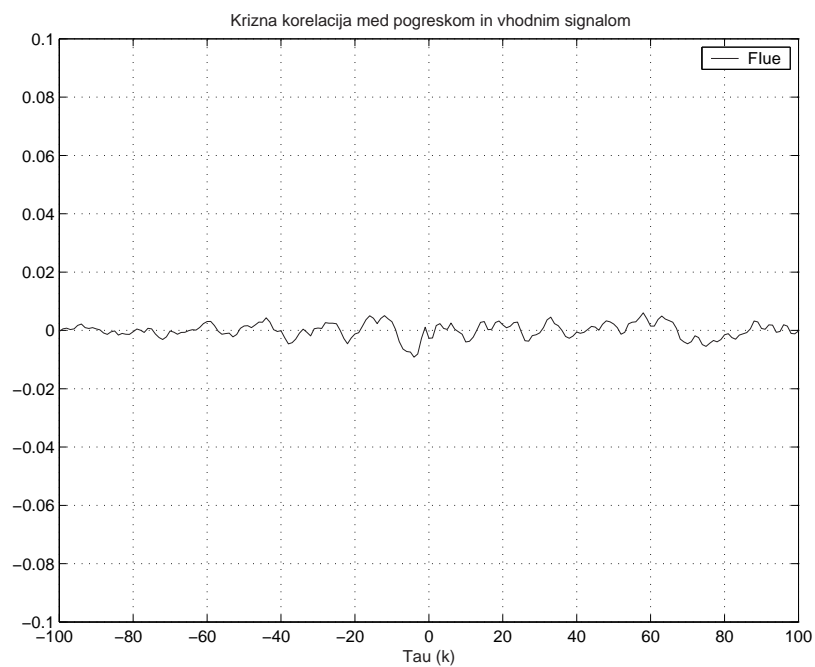
Slika 4.11: Simulacija z "naivno" in "eksaktno" metodo



Slika 4.12: Napaka simulacije GP pri "naivnem" in "eksaktnem" pristopu



Slika 4.13: Primerjava variance "naivne" in "eksaktne" metode na odseku signala, $u(k)$ povečan zaradi ilustrativnosti

Slika 4.14: Avtokorelacija napake $\Phi_{ee}(\tau)$ Slika 4.15: Križna korelacija med vzbujevalnim vhodnim signalom in napako $\Phi_{ue}(\tau)$

Če primerjamo enokoračno predikcijo GP modela z vključenimi lokalnimi modeli in GP model, ki vsebuje samo funkcijske vrednosti (sliki 3.5 in 4.6 oz. sliki 3.6 in 4.7) opazimo, da se modela v ožjem območju okoli ravnotežne krivulje (sistem se ponavadi večino časa nahaja na ali ob ravnotežni krivulji) bistveno ne razlikujeta. A GP model na osnovi funkcijskih vrednosti je bil naučen z $N = 200$ vhodno-izhodnimi pari, LM GP model pa z $N = 12 + 2 \cdot 6 = 24$ vhodno izhodnimi pari.

4.7.2 Dinamični sistem drugega reda

V tem poglavju bomo vrednotili pristop s primerom modela z vključenimi lokalnimi modeli za sistem drugega reda, opisan z enačbo (3.3).

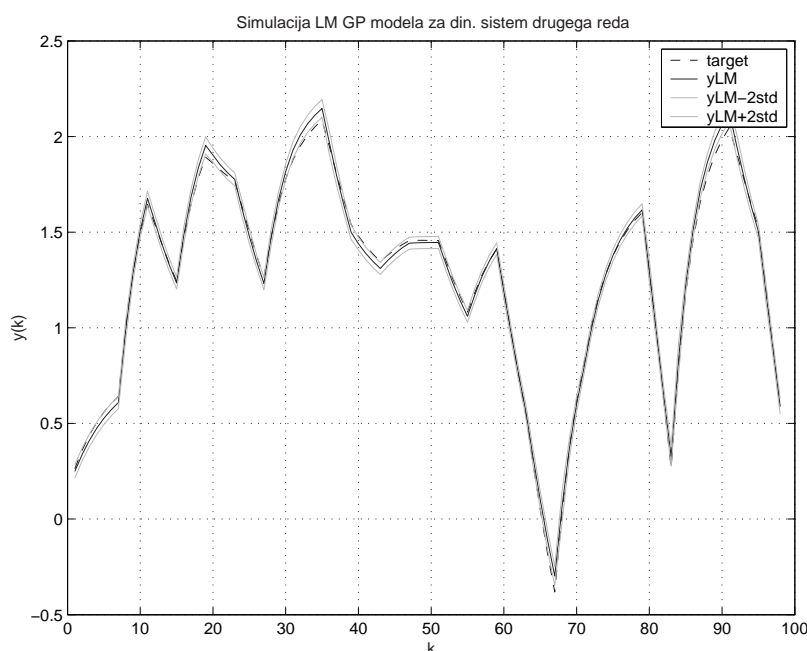
Izbrali smo področje delovanja med $u_{min} = -2$ in $u_{max} = 4$. Za učenje in vrednotenje sta bila uporabljena ista signala kot pri GP modelu brez vključenih lokalnih modelov (razdelek 3.3.2, sliki 3.11 in 3.12).

V GP model smo vključili pet lokalnih modelov, njihovi parametri so podani v tabeli 4.2. Velikostni red variance lokalnih modelov je 10^{-2} . Trem lokalnim modelom smo dodali 16 učnih točk, ki niso ležale na ravnotežni krivulji. Učni vektor je sestavljalo skupaj 41 točk.

Tabela 4.2: Parametri uporabljenih lokalnih modelov za nelinearni dinamični sistem drugega reda

U_{eq}	Y_{eq}	$-a_1$	$-a_2$	b_1	b_2
-1.3741	-1.5000	0.8504	-0.0500	0.2320	0
-0.4822	-0.5000	0.8800	-0.0500	0.1820	0
0.8711	0.8000	0.9088	-0.0500	0.1170	0
2.0325	1.6000	0.9101	-0.0500	0.0770	0
2.9139	2.0000	0.8957	-0.0500	0.0570	0

Rezultat simulacije identificiranega GP modela z vključenimi lokalnimi modeli je prikazan na sliki 4.16. Na sliki 4.17 je prikazana absolutna napaka simulacije $|e(k)| = |y_{GP}(k) - y(k)|$ skupaj z napovedano dvakratno standardno deviacijo 2σ . Kot pri GP modelu brez vključenih lokalnih modelov tudi tu velja, da model dobro identificira sistem. Rezultati so boljši od rezultatov, dobljenih z GP modelom,



Slika 4.16: Simulacija GP modela za nelinearni dinamični sistem drugega reda

naučenim s 83 točkami, kljub temu, da je bil LM GP model učen s približno dvakrat manj učnimi točkami.

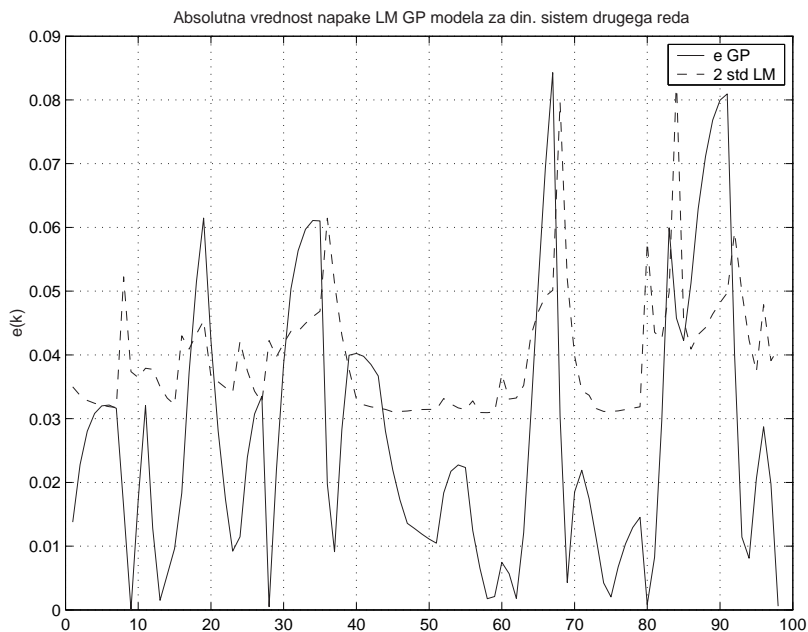
Sliki 4.18 in 4.19 prikazujeta avtokorelacijo napake simulacije LM GP modela $e(k)$ in križno korelacijo med napako identifikacije $e(k)$ ter vhomom v sistem $u(k)$.

Vrednosti cenilk za odziv LM GP modela sta $SE = 0.0012$ in $LD = -1.7241$.

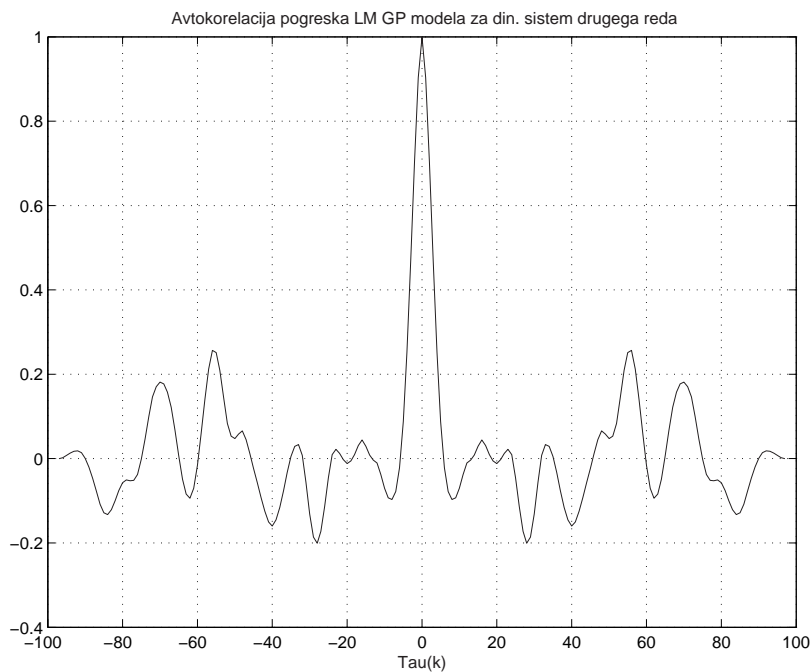
Tudi ta sistem smo poizkusili simulirati po "eksaktni" metodi. Za njeno delovanje je potrebno v vsakem koraku izračunavati križne kovariance med (napovedanimi) preteklimi izhodi (3.24). Vrednotenje je pokazalo, da problem računanja križnih kovarianc ni tako preprost, kot se je zdelo in da simulacija za sisteme višjega reda ne daje pričakovanih rezultatov.

Povzemimo glavne prednosti GP modela z vključenimi lokalnimi modeli pred navadnim GP modelom:

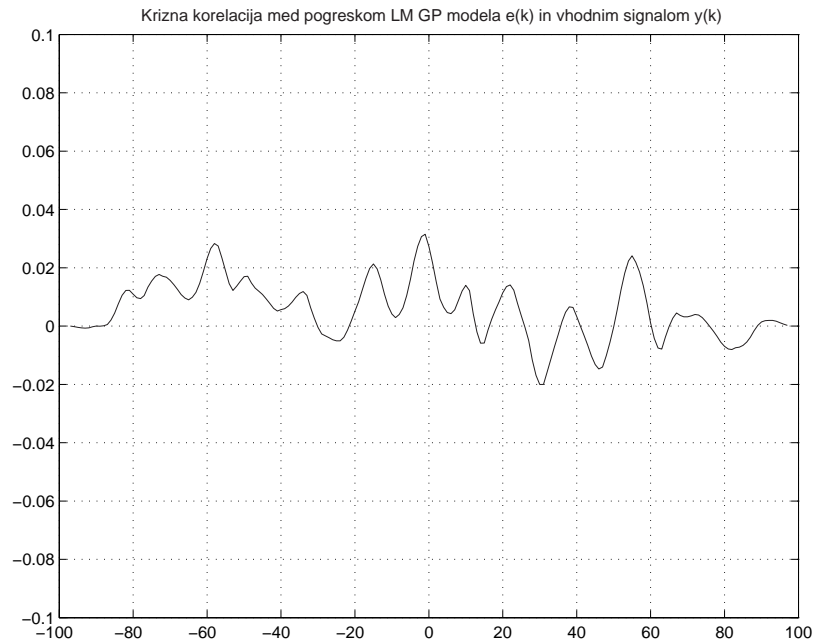
- Omogoča elegantno vključitev predhodnega znanja o nekem delu sistema. V območju ravnotežne krivulje namreč ni težko priti do zadostne količine podatkov za izgradnjo lokalnih linearnih modelov.
- Izven ravnotežne krivulje ni treba graditi lokalnih modelov, kar je ponavadi



Slika 4.17: Absolutna napaka $|e(k)|$ simulacije GP modela za nelinearni dinamični sistem drugega reda



Slika 4.18: Avtokorelacija napake predikcije GP modela Φ_{ee} za nelinearni dinamični sistem drugega reda



Slika 4.19: Križna korelacija vhoda in napake predikcije GP modela Φ_{ue} za nelinearni dinamični sistem drugega reda

zelo oteženo zaradi majhnega števila podatkov, ki jih lahko pridobimo z meritvami. S tem ima LM GP model prednost pred modelom sestavljenim iz mreže lokalnih modelov.

- Zaradi manjšega števila točk (v primerjavi z GP modelom na osnovi funkcijskih vrednosti) se zmanjša kovariančna matrika \mathbf{K} in posledično lahko zelo poveča hitrost učenja. Za koliko se zmanjša obseg modela in hitrost učenja pa je odvisno od nelinearnosti sistemov, ki jih modeliramo.

Do sedaj smo za ponazoritve procesov uporabljali le sisteme opisane z nelinearnimi diferenčnimi enačbami. To smo storili namenoma, da bi se lahko osredotočili na specifične lastnosti pristopa k identifikaciji dinamičnih sistemov z GP.

Navkljub temu bi radi opozorili na določene probleme, ki jih do sedaj nismo izpostavljali.

- Izbira reda identificiranega modela sistema ni nujno enaka redu nelinearnega sistema, ki ga identificiramo (če ga sploh vemo), ampak je lahko precej večji

[11].

- Procesi, ki v določenih področjih delovanja pri istih vhodnih vrednostih dajejo različne izhodne vrednosti (npr. histereza) lahko predstavljajo za identifikacijo problem. GP pri takem procesu napovejo povprečno vrednost in primerno razširijo varianco.

Kadar identificiramo model na osnovi meritev moramo paziti tudi na čas vzorčenja in kvaliteto informacije v identifikacijskih signalih, ki je lahko omejena tudi s področjem obratovanja naprave. Primer identifikacije z GP na osnovi meritev prikazuje naslednje poglavje.

5. Primer identifikacije modelne naprave z GP

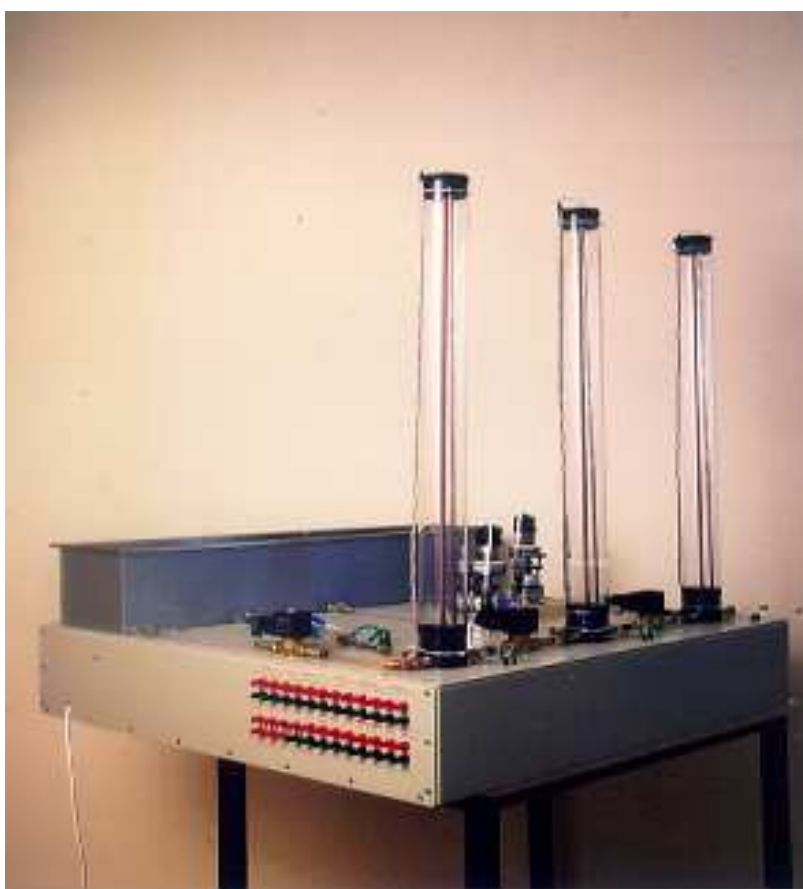
Prava vrednost identifikacijskih metod se pokaže šele pri njihovi uporabi na signalih, dobljenih z meritvami. V tem poglavju bo predstavljena identifikacija laboratorijske modelne naprave treh tankov, ki je prikazana na sliki 5.1.

Sistem zaradi blage nelinearnosti morda ni najbolj ilustrativen za prikaz delovanja LM GP modelov za identifikacijo, vendar smo ga uporabili zaradi dobrega poznavanja procesa na Odseku za sisteme in vodenje Inštituta Jožef Stefan, kjer smo identifikacijo izvajali. Poleg tega se je izkazalo, da v našem dosegu ni na voljo večjega števila stabilnih naprav z izraženo nelinearnostjo, na katerih bi lahko izražali potrebne meritve.

5.1 Opis naprave

Laboratorijska modelna naprava treh tankov je nelinearni dinamični sistem namenjen za preizkušanje metod modeliranja in vodenja [1, 20] na Odseku za sisteme in vodenje na Inštitutu Jožef Stefan. Procesna shema modelne naprave je predstavljena na sliki 5.2.

Modelna naprava je sestavljena iz treh tankov $R1$, $R2$ in $R3$, ki so prek cevi povezani s shranjevalnikom tekočin $R0$. Med sabo so paroma povezani tanki $R1$ in $R2$ ter $R2$ in $R3$. Tanka $R1$ in $R3$ imata za črpanje iz shranjevalnika v ceveh vgrajeni vsaka svojo črpalko ($P1$ in $P2$), ki ju poganjata enosmerna motorja s trajnim magnetom. Hitrost vrtenja motorjev je regulirana z analognim regulatorjem. Časovna konstanta spremembe hitrosti vrtenja motorjev je zelo majhna v primerjavi z dinamiko višine tekočin v treh tankih in jo lahko zanemarimo (predpostavljamo, da ni zakasnitve med referenčno in pravo hitrostjo). Napravo sestavlja še pet ventilov:

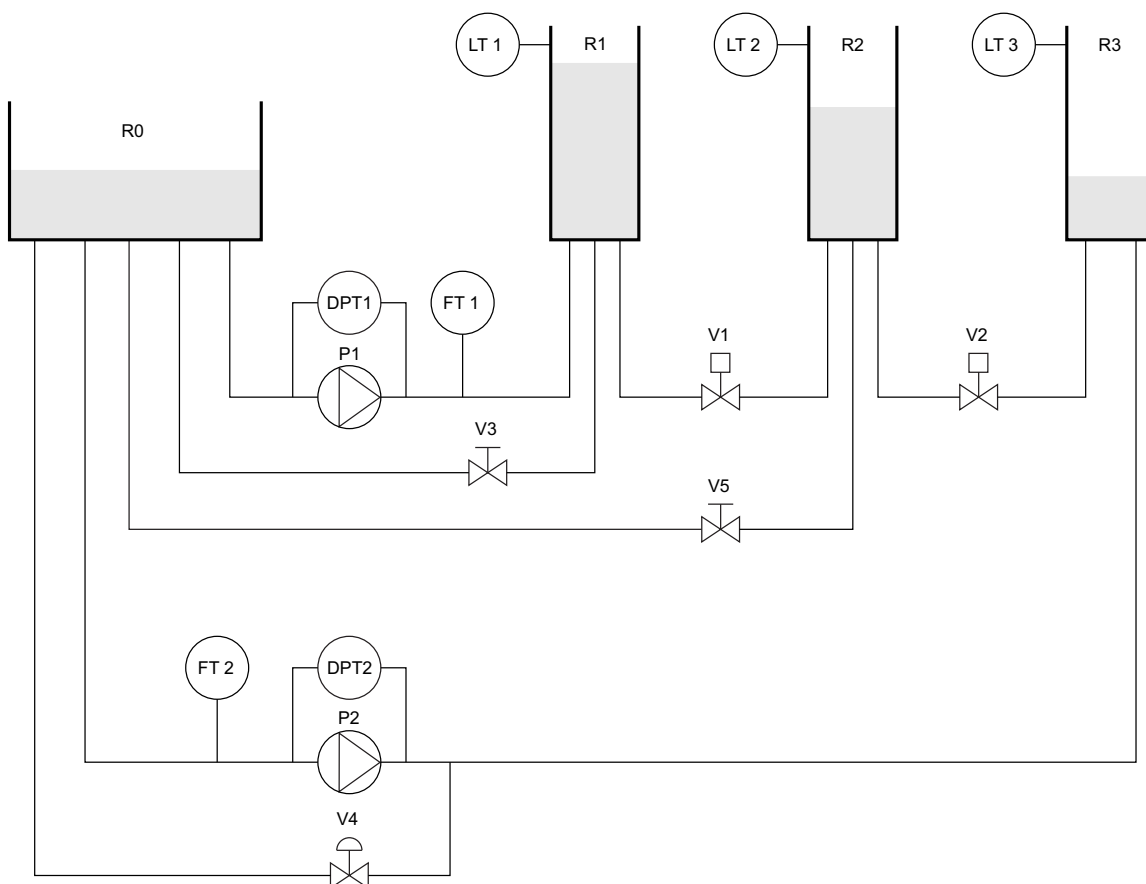


Slika 5.1: Laboratorijska modelna naprava treh tankov

- regulacijski ventil V_4 , ki ga poganja enosmerni motor,
- dva dvopoložajna ventila V_1 in V_2 ,
- dva ročna ventila V_3 in V_5 .

Presek shranjevalnika R_0 je precej večji od presekov tankov R_1 , R_2 ter R_3 , zato lahko nivo tekočine v njem smatramo praktično za konstanten.

Na napravo vplivamo preko hitrosti vrtenja obeh črpalk in odprtostjo ventilov. Analogno digitalni (A/D) in digitalno analogni (D/A) vmesnik *Burr-Brown PCI 20000 system* povezuje vhode in izhode naprave z okoljem *Matlab Simulink*.



Slika 5.2: Procesna shema modelne naprave treh tankov

5.1.1 Seznam senzorjev

Seznam senzorjev naprave je podan v tabeli 5.1. Nekatere vrednosti izhodov senzorjev zaradi lažje preglednosti niso izražene z mednarodnim sistemom enot (SI). To velja za spremembe tlaka (cmH_2O), velikosti pretokov (cm^3/s) ter nivoje tekočine v posodah (cm).

Zaradi konstrukcije črpalk ni mogoče pridobiti podatka o njeni hitrosti, zato je hitrost črpalke namesto v $[s^{-1}]$ izražena z $[V]$ – uporabljen bo signal na ustreznem A/D kanalu. Podobno velja tudi za ventile, kjer je njihov položaj izražen v $[V]$. Vse veličine modela so predstavljene z napetostnim signalom. Meritve nivojev se opravljajo s kapacitivnimi senzorji. Tlačne razlike so merjene s *Sensortronics HCXM 100 D6V* senzorji. Oboji imajo linearno karakteristiko.

Pretoka sta merjena posredno preko padca tlaka na Venturijevih ceveh *VE1* in *VE2* z že prej omenjenimi senzorji. Zveza med pretokom skozi cev in izhodom sensorja je kvadratna.

Hitrost vrtenja črpalk *P1* in *P2* je merjena z induktivnim principom. Permanentni magnet na rotorju črpalke inducira napetost na tuljavi, pritrjeni na statičnem delu črpalke. Napetost se pretvori v pulzni signal, ta pa preko frekvenčno-napetostnega pretvornika v napetost na izhodu sensorja.

5.1.2 Seznam aktuatorjev

Na pretoke tekočine med posodami vplivamo prek hitrosti črpalk *P1* in *P2* ter položaji servo-ventila *V4*, dvopoložajnih ventilov *V1* in *V2* ter ročnih ventilov *V3* in *V5*. Natančneje so aktuatorji opisani v tabeli 5.2.

Črpalke vsebujejo povratnozančno regulacijo, ki je dovolj hitra, da lahko razliko med referenčno in pravo hitrostjo motorčka, ki jo poganja, v primerjavi z dinamiko spreminjanja višine tekočine v tankih smatramo za zanemarljivo.

5.1.3 Uporabljeni podsistem naprave

Naprava omogoča realizacijo sistema tretjega reda, vendar so meritve izvedljive v preozkem področju, zaradi česar nelinearnost ne pride do izraza. To bi zmanjšalo

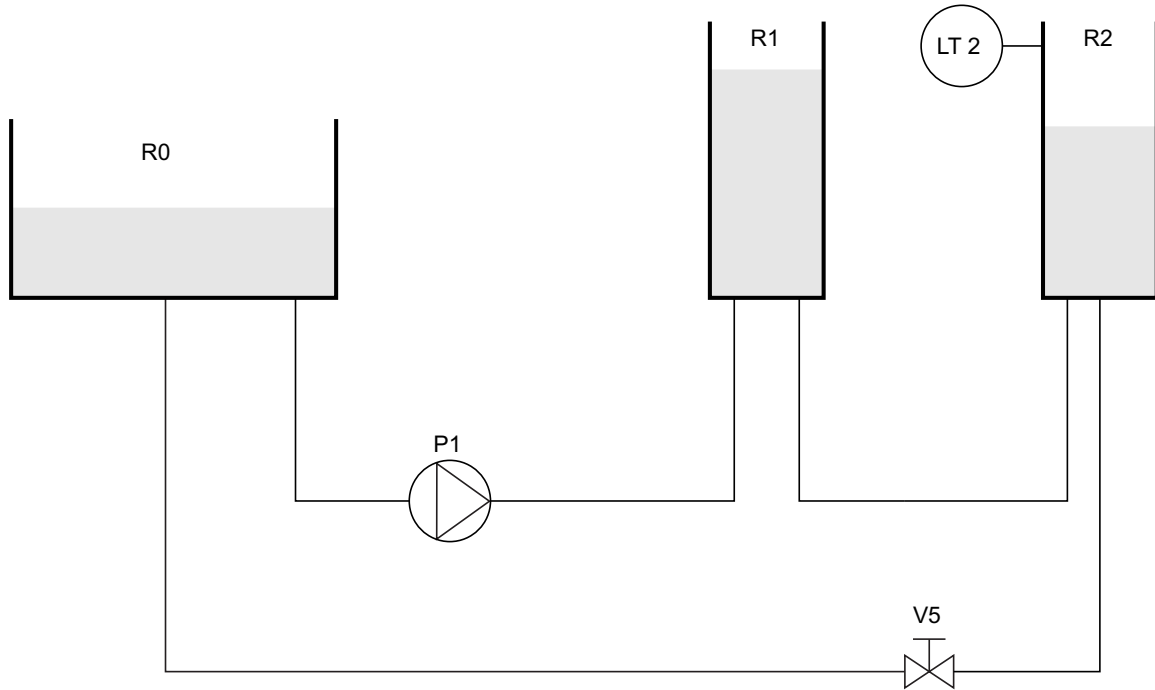
ilustrativnost primera. Zato sta bila uporabljena samo tanka $R1$ in $R2$, tank $R3$ pa je bil "odrezan" z zaprtjem ventila $V2$. Prav tako je bil zaprt ročni ventil $V3$. Vhod v sistem je predstavljala hitrost vrtenja ω_1 črpalke $P1$, ki črpa tekočino v rezervoar $R1$ od koder se pretaka naprej v tank $R2$. Iz tega se skozi (delno) odprti ročni ventil $V5$ tekočina pretaka nazaj v rezervoar $R0$. Izhod iz sistema predstavlja nivo tekočine h_2 tanka $R2$. Na sliki 5.3 je shema uporabljenega podsistema.

Tabela 5.1: Tabela spremenljivk modela

Simbol	Spremenljivka	Enota	Področje	Opis
LT1	h_1	cm	0 - 60	nivo tekočine v tanku $R1$
LT2	h_2	cm	0 - 60	nivo tekočine v tanku $R1$
LT3	h_3	cm	0 - 60	nivo tekočine v tanku $R1$
DPT1	p_1	$cmH20$	0 - 100	tlačna razlika na črpalke $P1$
DPT2	p_2	$cmH20$	0 - 100	tlačna razlika na črpalke $P2$
FT1	Q_1	cm^3/s	0 - 30	pretok skozi črpalke $P1$
FT2	Q_2	cm^3/s	0 - 30	pretok skozi črpalke $P2$
-	1	-	0 - 6	hitrost vrtenja črpalke $P1$
-	2	-	0 - 6	hitrost vrtenja črpalke $P2$
-	s_4	-	0 - 10	položaj regulacijskega ventila $V4$
-	U_1	V	0 - 10	napetost na motorju črpalke $P1$
-	I_1	-	0 - 0.3	tok skozi motor črpalke $P1$
-	U_2	V	0 - 10	napetost na motorju črpalke $P2$
-	I_2	-	0 - 0.3	tok skozi motor črpalke $P2$

Tabela 5.2: Tabela aktuatorjev modela

Simbol	Sprem.	Interval [V]	Opis
P1	ω_{1r}	0 - 6	referenca regulatorja hitrosti motorja črpalke $P1$
P2	ω_{2r}	0 - 6	referenca regulatorja hitrosti motorja črpalke $P2$
-	U4	0 - 10	referenca položaja regulacijskega ventila $V4$



Slika 5.3: Procesna shema uporabljenega podsistema

5.1.4 Matematični model naprave

Za modeliranje je bila naprava najprej razdeljena na posamezne elemente. Za vsak element posebej je bil nato razvit fizikalni oz. polfizikalni model. Razvoj matematičnega modela naprave je podrobno opisan v [1].

Del naprave, ki nas zanima, lahko predstavimo z dvema stanjema v prostoru stanj: nivojem tekočine v tanku $R1$ h_1 in nivojem tekočine v tanku $R2$ h_2 . Nivo h_2 je hkrati tudi izhod sistema, vhod pa je že omenjena hitrost ω_1 črpalke $P1$. V matematični obliki lahko sistem predstavimo:

$$\dot{h}_1 = \frac{1}{A_1} \left(Q_1(\omega_1, h_1) - Q_{12}(h_1, h_2) \right) \quad (5.1)$$

$$\dot{h}_2 = \frac{1}{A_2} \left(Q_{12}(h_1, h_2) - Q_2(h_2) \right) \quad (5.2)$$

$$Q_1(h_1, \omega_1) = k_{1F}(h_{1stat}(\omega_1) - h_1) + k_{2F}\sqrt{h_{1stat}(\omega_1) - h_1} \quad (5.3)$$

$$Q_{12}(h_1, h_2) = k_{V1}\sqrt{\rho g(h_1 - h_2)} \quad (5.4)$$

$$Q_2(h_2) = k_{V5}\sqrt{\rho g h_2} \quad (5.5)$$

$$k_{1F}(\omega_1) = k_{1F0} + k_{1F1}\omega_1 + k_{1F2}\omega_1^2 \quad (5.6)$$

$$k_{2F}(\omega_1) = k_{2F0} + k_{2F1}\omega_1 \quad (5.7)$$

kjer sta A_1 in A_2 preseka tankov $R1$ in $R2$, k_{V1} konstanta ventila $V1$, k_{V5} konstanta ventila $V5$, k_{1F} in k_{2F} parametra črpalke $P1$, $h_{1stat}(\omega_1)$ pa ustaljena višina h_1 pri neki hitrosti vrtenja črpalke ω_1 (statična karakteristika (ω_1, h_{1stat})). Vrednosti parametrov so podane v tabeli 5.3.

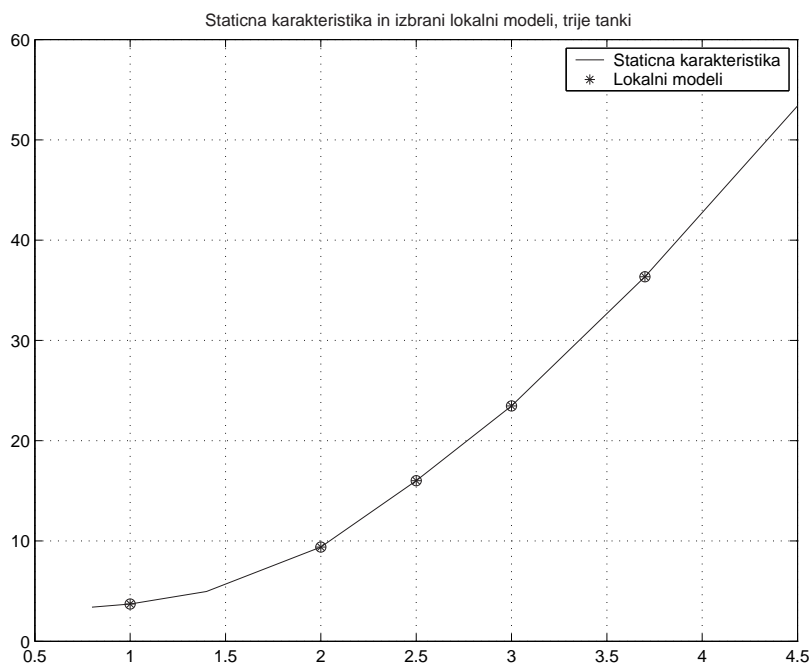
Tabela 5.3: Tabela konstantnih parametrov modela

Konstanta	Vrednost	Enota	Opis
A_1	19.16	cm^2	preseki tanka $R1$
A_2	19.16	cm^2	preseki tanka $R2$
k_{1F0}	6.595e-1		koeficient v izrazu za k_{1F}
k_{1F1}	-2.128e-1		koeficient v izrazu za k_{1F}
k_{1F2}	1.751e-2		koeficient v izrazu za k_{1F}
k_{2F0}	9.352e-1		koeficient v izrazu za k_{2F}
k_{2F1}	3.024e-1		koeficient v izrazu za k_{2F}
k_{V1}	2.72		tokovni koeficient v smeri od $R1$ do $R2$
k_{V1B}	2.31		tokovni koeficient v smeri od $R2$ do $R1$
k_{V5}	ni pod.		tokovni koeficient v smeri od $R2$ do $R0$

5.2 Opis področja delovanja naprave in meritev

Edini vhod v izbrani sistem drugega reda predstavlja napetost $U[V]$, ki krmili hitrost ω_1 črpalke $P1$. Edini izhod iz naprave je višina stolpca tekočine $h_2[cm]$, merjena s kapacitivnim senzorjem in pretvorjena iz $[V]$ v $[cm]$. Za poenostavitev bomo vrednosti od sedaj navajali brez merskih enot. Na sliki 5.4 je prikazana statična karakteristika sistema. Področje delovanje omejujejo višine tankov, saj jih nivoji tekočine ne morejo preseči. Področje delovanje je bilo s preizkušanjem izbrano tako, da črpalka $P1$ nikoli ni ustvarila dovolj pritiska, da bi se nivo tekočine v prvi posodi vzdignil previsoko ($U_{max} = 4$).

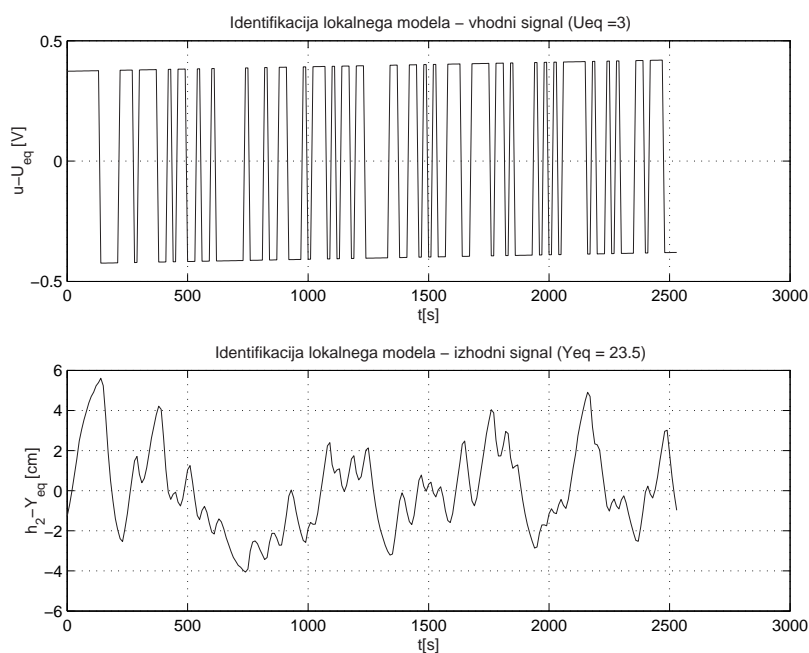
Merjenje je potekalo na dva načina:



Slika 5.4: Statična karakteristika sistema

- merjenje za pridobitev lokalnih modelov v različnih točkah na statični karakteristiki,
- merjenje, kjer smo s spremembami vhoda vplivali na sistem, da bi dobili podatke izven ravnotežne krivulje.

Za pridobitev lokalnega modela v neki ravnotežni točki moramo sistem s statičnim vhodnim signalom U_{eqi} najprej pripeljati v to točko in ga tam vzbujati s signalom, ki je frekvenčno dovolj bogat, da lahko iz meritev izluščimo dovolj informacije. Z iterativnim postopkom identifikacij pri različnih časih vzorčenja smo ugotovili, da je čas vzorčenja $T_s = 10s$ primeren za naš sistem. Ko smo sistem pripeljali v delovno točko, smo osnovnemu vhodnemu signalu ($U = U_{eqi}$) dodali psevdo-naključni binaren signal [14], kjer se najhitrejša sprememba (takt signala) zgodi v $2T_s = 20s$. Psevdo-naključni binarni signal je imel amplitudo $\Delta U = 0.4$. Ta amplituda določa področje veljavnosti modela. Linearni model je bil v ravnotežni točki identificiran z algoritmom pomožnih spremenljivk programskega paketa Matlab (IV_4). Na sliki 5.5 sta prikazana vhodni in izhodni signal za točko $U_{eq} = 3, Y_{eq} = 23.5$ in $\Delta U = 0.4$. Enako, le z dvakrat hitrejšim taktom signala $T_s = 10s$, je bil dobljen signal za vrednotenje dobljenih linear-



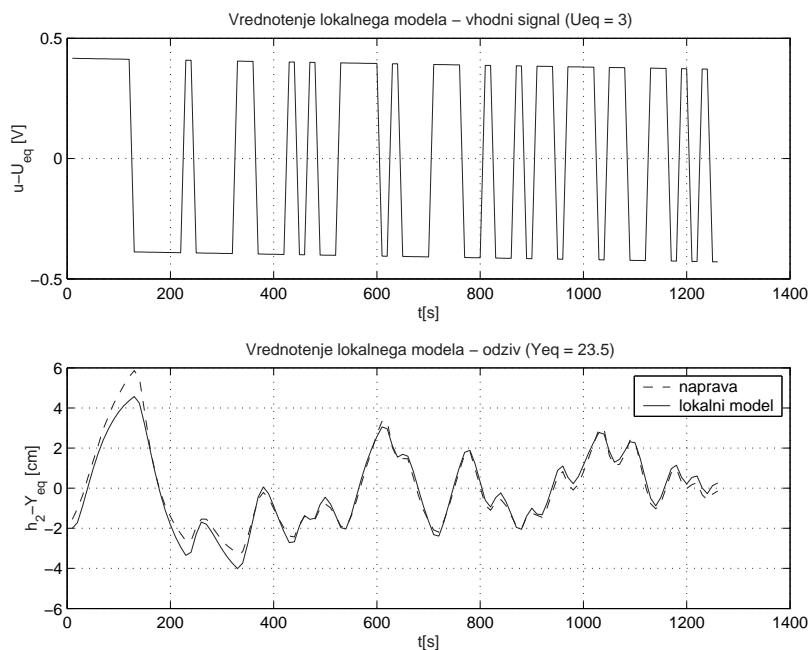
Slika 5.5: Signal za identifikacijo lokalnega modela v delovni točki $(U_{eq}, Y_{eq}) = (3, 23.5)$

nih modelov. Na sliki 5.6 je prikazano vrednotenje linearnega modela v točki $U_{eq} = 3, Y_{eq} = 23.5$.

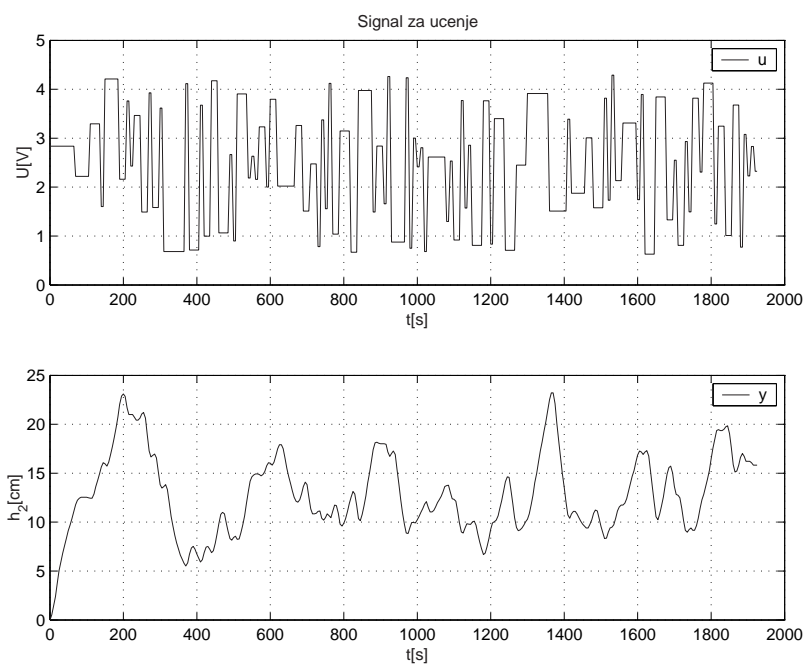
Z drugim tipom meritve smo pridobili učne podatke za območje izven ravnotežne krivulje in podatke za vrednotenje GP modela. Vzbujevalni signali so bili realizirani podobno kot prej omenjeni psevdo-naključni binarni signal, le da je vrednost signala ob spremembi lahko zavzela katerokoli vrednost med $U_{min} = 0.8$ in $U_{max} = 4$. Merjenje je potekalo z dvema vhodnima signaloma z različnim taktom spreminjanja amplitude. Najhitrejša možna sprememba prvega signala je bila $T_s = 10s$, drugega pa $2T_s = 20s$. Prvi vhodni in prvi izhodni signal, uporabljena za učenje, sta prikazana na sliki 5.7. Drugi vhodni in drugi izhodni signal, uporabljena za vrednotenje, sta prikazana na sliki 5.8.

5.3 Identifikacija GP modela

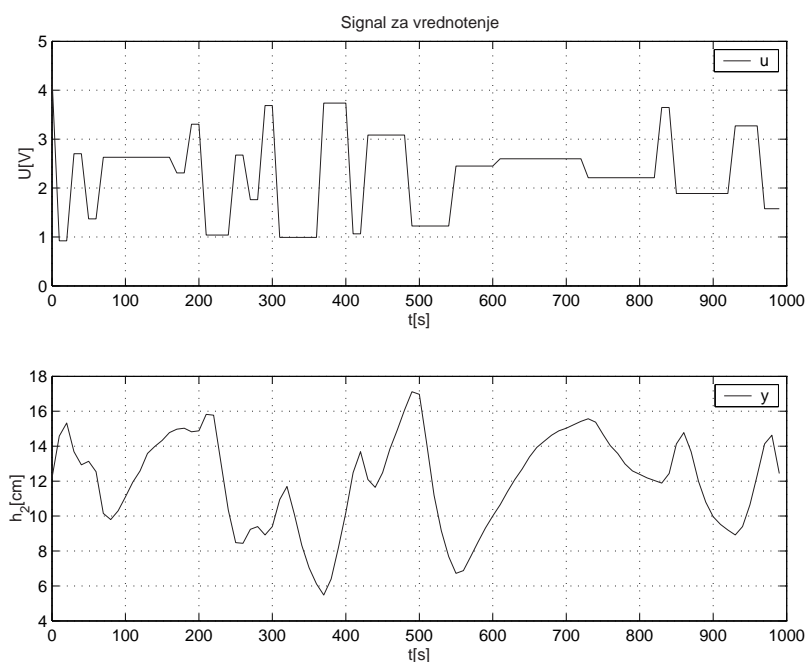
Za učenje GP modela drugega reda je bilo uporabljenih 162 vhodno-izhodnih podatkov iz signala za učenje na sliki 5.7. Rezultat simulacije GP modela je prikazan na sliki 5.9, absolutna vrednost napake simulacije pa skupaj s standardno



Slika 5.6: Signal za vrednotenje lokalnega modela v delovni točki $(U_{eq}, Y_{eq}) = (3, 23.5)$



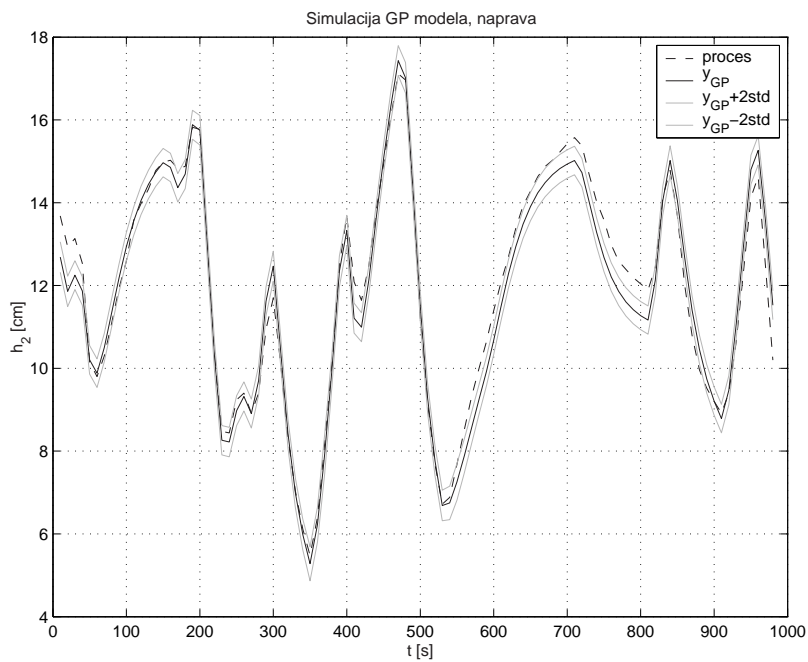
Slika 5.7: Signal za učenje, najhitrejša sprememba v času $T = 10s$



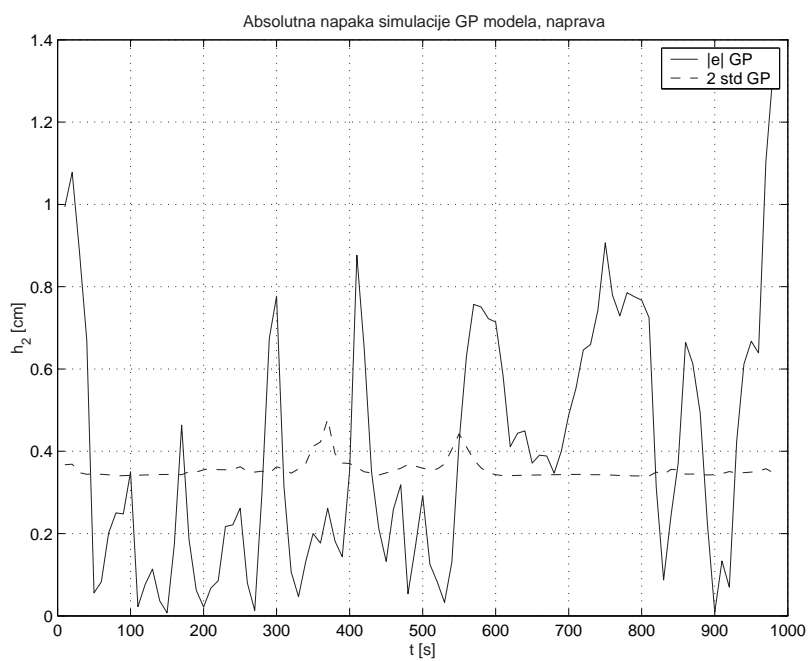
Slika 5.8: Signal za vrednotenje, najhitrejša sprememba v času $T = 20s$

deviacijo na sliki 5.10. Na slikah 5.11 sta prikazani avtokorelacija napake GP modela in križna korelacija med napako GP modela in vhodnim signalom v GP model u . Delovanje "eksaktnega" GP modela s porazdelitvijo kot vhodom je prikazano na slikah 5.13 in 5.14. Vrednosti cenilk odziva "naivnega" modela sta $SE = 0.2439$ in $LD = 3.1343$, vrednosti cenilk izhoda "eksaktnega" modela $SE = 0.2441$ in $LD = 0.8978$.

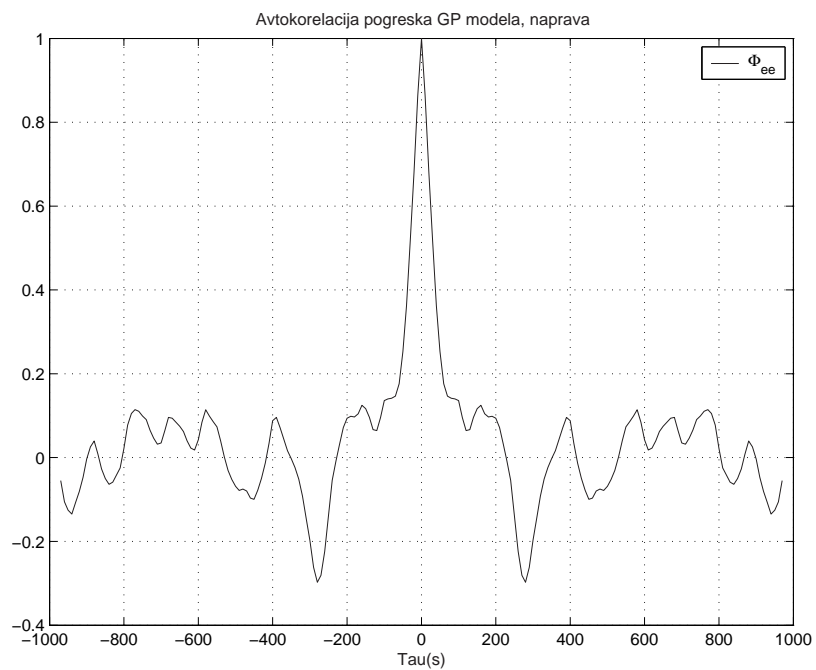
Oba modela zadovoljivo ujameta delovanje sistema, stopnja zaupanja v predikcijo "eksaktnega" modela pa bolje odseva dejansko stanje kot tista "naivnega" modela.



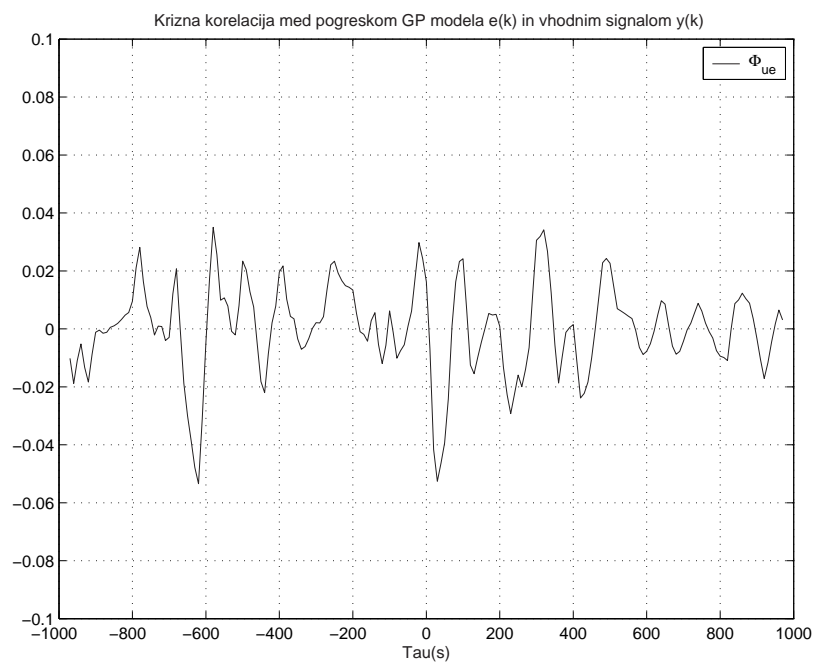
Slika 5.9: Simulacija GP modela



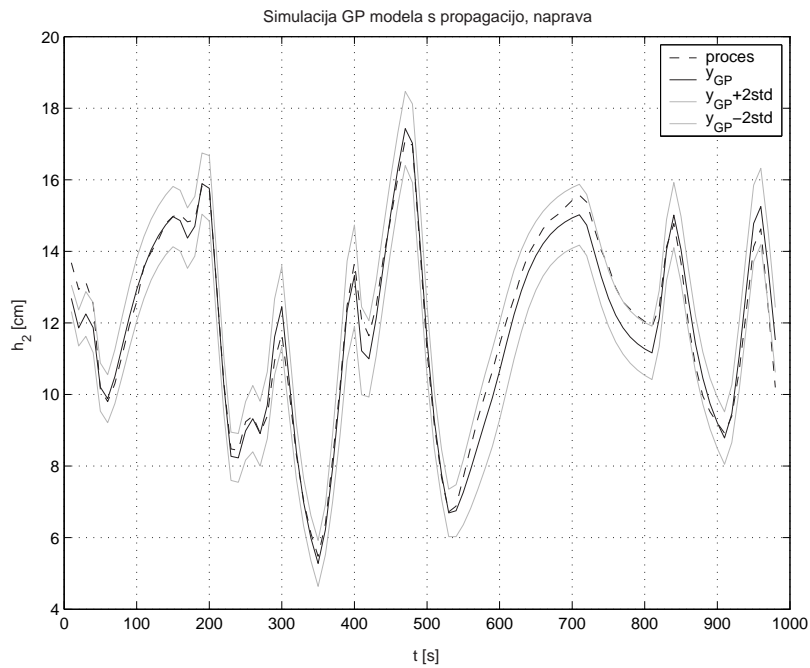
Slika 5.10: Napaka simulacije GP modela



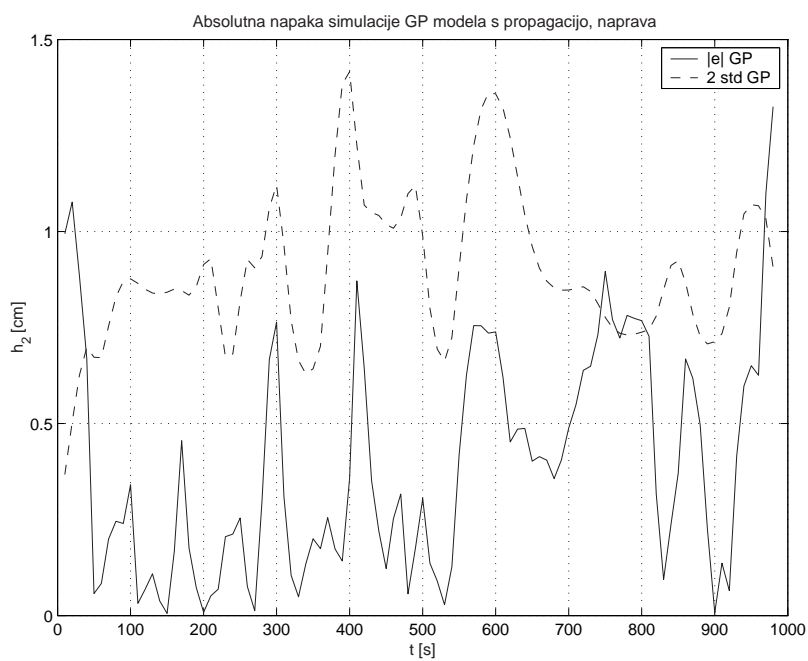
Slika 5.11: Avtokorelacija napake GP modela



Slika 5.12: Križna korelacija napake GP modela in vhodnega signala u



Slika 5.13: Simulacija "eksaktnega" GP modela



Slika 5.14: Napaka simulacije "eksaktnega" GP modela

5.4 Identifikacija GP modela z vključenimi lokalnimi modeli

GP model z vključenimi lokalnimi modeli (LM GP model) drugega reda smo učili s petimi lokalnimi modeli in enajstimi točkami izven ravnotežne krivulje, kar skupaj sestavlja 36 vhodno-izhodno učnih parov. Središča lokalnih modelov so prikazani na statični karakteristiki na sliki 5.4.

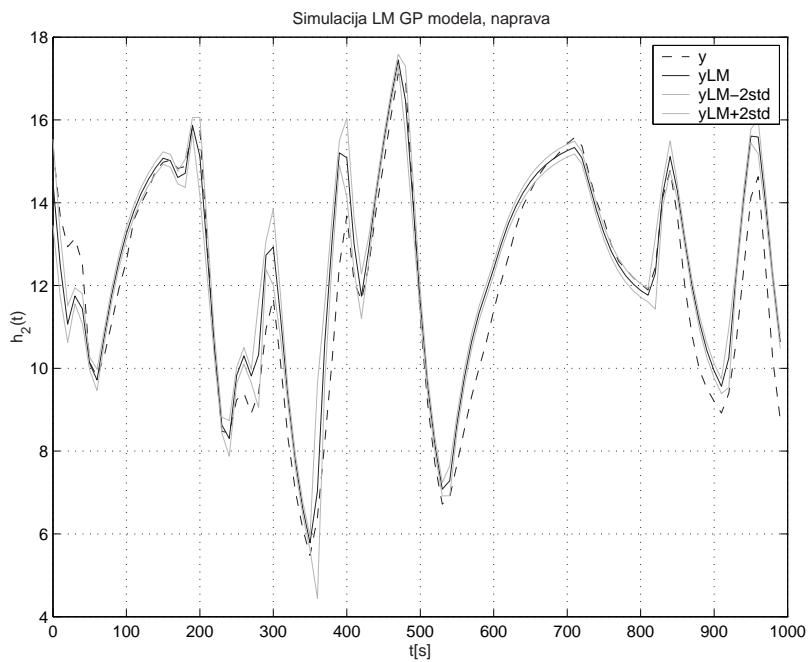
Parametri modelov so podani v tabeli 5.4. Diagonalni elementi kovariančna matrike so bili reda velikosti 10^{-3} do 10^{-4} . Linearni model, ki opisuje obnašanje modela v delovni točki (U_{eq}, Y_{eq}) opisuje enačba:

$$y(k) = -a_1y(k-1) - a_2y(k-2) + b_1u(k-1) + b_2u(k-2) \quad (5.8)$$

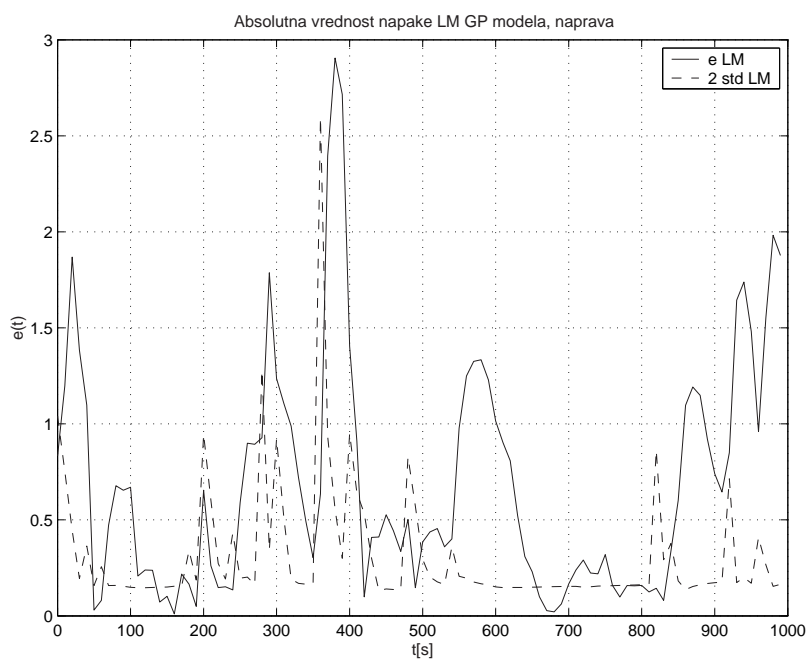
Tabela 5.4: Parametri uporabljenih lokalnih modelov za model treh tankov

U_{eq}	Y_{eq}	$-a_1$	$-a_2$	b_1	b_2
1.0000	3.7136	0.7573	-0.0580	0.2903	0.0704
2.0000	9.3872	0.9328	-0.1214	0.9206	0.8613
2.5000	16.0096	1.1594	-0.2912	0.7792	0.8767
3.0000	23.4668	1.1674	-0.2799	0.9212	1.0892
3.7000	36.3530	1.3712	-0.4446	0.6829	0.7032

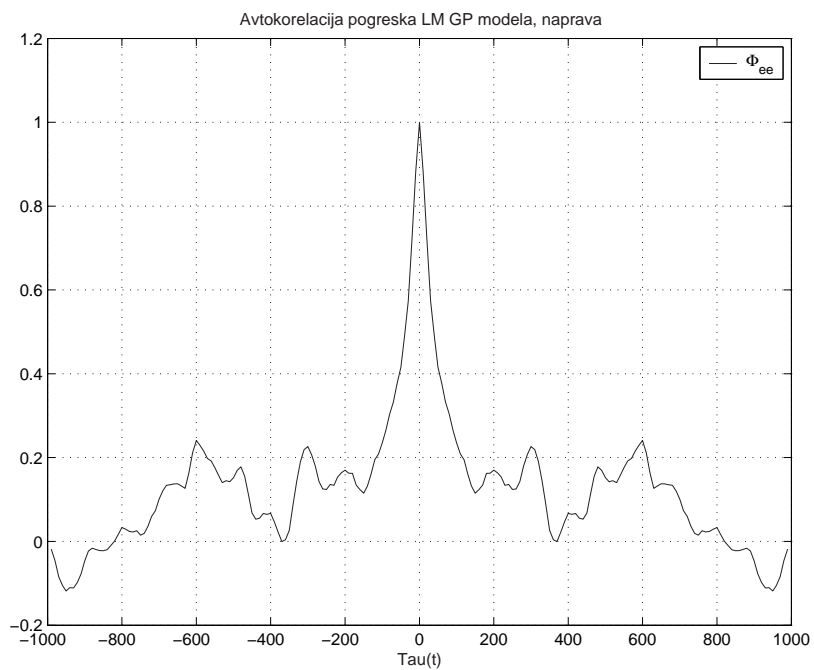
Rezultat simulacije GP modela z vključenimi lokalnimi modeli je prikazan na sliki 5.15, absolutna vrednost napake simulacije pa skupaj s standardno deviacijo na sliki 5.16. Avtokorelacija napake predikcije LM GP modela je prikazana na sliki 5.17, križna korelacija napake predikcije in vhoda $u(t)$ pa na sliki 5.18. Cenilki za prikazani odziv LM GP modela imata vrednosti $SE = 0.8518$ in $LD = 32.98$. Zaradi že v poglavju 4.5 omenjenih problemov s propagacijo variance pri GP modelih z vključenimi lokalnimi modeli sistema nismo identificirali z "eksaktnim" GP modelom.



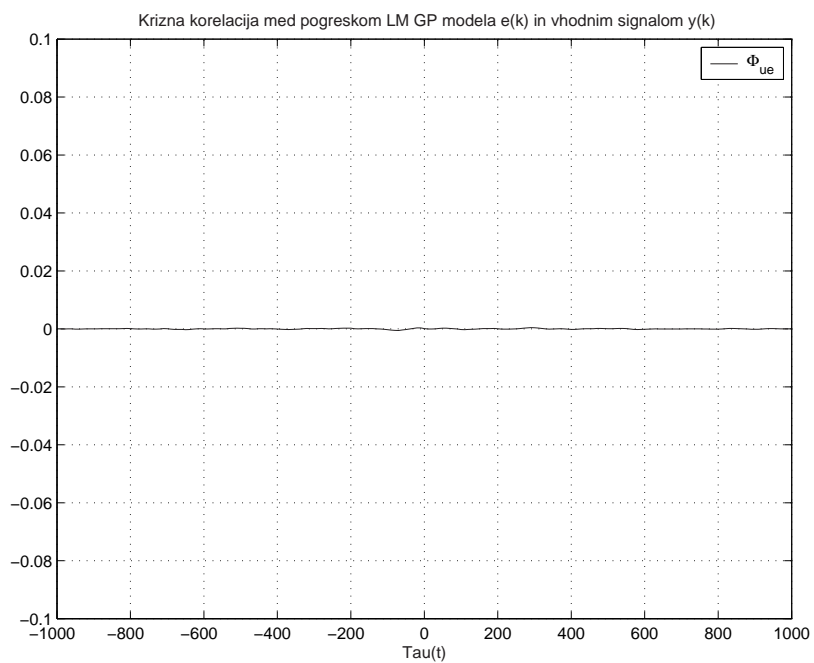
Slika 5.15: Simulacija GP modela



Slika 5.16: Napaka simulacije GP modela



Slika 5.17: Avtokorelacija pogreška GP modela z vključenimi LM



Slika 5.18: Križna korelacija pogreška simulacije GP modela z vključenimi LM

5.5 Primerjava rezultatov in komentar

Rezultati identifikacije niso slabi, čeprav odziva ne ujamejo povsem. Tako GP model kot GP model z vključenimi lokalnimi modeli ujameta glavne značilnosti sistema. Odziv GP modela z vključenimi lokalnimi modeli je nekoliko slabši. Odziv bi lahko izboljšali s povečanjem števila učnih točk, predvsem tistih stran od ravnotežne krivulje. Rezultati pa vseeno dajejo vtis, da bi GP modeli (navadni in z vključenimi lokalnimi modeli) lahko bili uporabni pri identifikaciji nelinearnih sistemov.

6. Zaključek

V magistrskem delu je bil predstavljen osnovni princip identifikacije dinamičnih sistemov z Gaussovimi procesi. Metoda spada med neparametrične identifikacijske metode in je dopolnilo identifikacije z umetnimi nevronskimi mrežami ter mehko logiko. Izhod iz Gaussovega procesa je predikcija normalne porazdelitve, opisana s srednjo vrednostjo in varianco. Predstavili smo tudi razširitev metode s porazdelitvijo kot vhodom v GP.

Vrednotenje rezultatov identifikacije z GP je bilo zadovoljivo. Predikcija GP modela je bila vseskozi blizu pravi vrednosti ali pa nas je na poslabšanje napovedi (zaradi pobega iz učnega področja) opozorila povečana varianca.

Opisana je bila tudi razširitev GPov z lokalnimi modeli, ki imajo nekaj zanimivih lastnosti. Tu bi lahko izpostavili uporabnost podatkov meritev na realnih procesih, ki jih je ponavadi veliko prevsem okoli ravnotežnih vrednosti, kot tudi potencial, da bo z metodo zaradi manjšega števila podatkov identifikacija hitrejša kot za GP model, ki vsebuje samo funkcijske vrednosti.

Metoda je zadovoljivo delovala pri identifikaciji umetnega dinamičnega sistema prvega reda in dinamičnega sistema drugega reda. Sistemov višjega reda od drugega zaradi obsežnosti GP modelov nismo uporabili za vrednotenje. Za uspešno se je izkazala propagacija celotne porazdelitve nazaj na vhod, kar je pomembno za simulacijo GP modelov. Vključevanje lokalnih modelov je bilo zaradi različne programske opreme za njihovo izvedbo težavnejše. Izvedba GP modelov z vključenimi lokalnimi modeli in propagacijo porazdelitve je uspela samo za dinamične sisteme prvega reda, realizacija za sisteme višjega reda predstavlja izziv za v bodoče.

Model laboratorijske modelne naprave je bil uspešno identificiran z GP modelom in GP modelom z vključenimi lokalnimi modeli, a velja opozoriti, da pri

simulaciji naprave z GP modelom z vključenimi lokalnimi modeli nismo uporabili razširjanja variance.

Metoda ima nekaj potencialnih prednosti pred podobnimi metodami za identifikacijo (Takagi-Sugeno mehki modeli, umetne nevronske mreže idr.). Glavna prednost in razlika je, da poleg napovedi daje tudi zaupanje v to napoved. Druga razlika je, da je potrebno optimizirati precej manjše število parametrov kot npr. pri nevronskih mrežah, kar pomeni lažjo optimizacijo z boljšimi možnostmi, da konča v globalnem minimumu. Z vključevanjem linearnih lokalnih modelov v metodo identifikacije z Gaussovimi procesi lahko nadomestimo veliko število sicer potrebnih učnih točk in tako zmanjšamo računsko breme. Verjetno še pomembnejša prednost vključevanja lokalnih modelov je možnost vključitve lokalne informacije v globalno ne da bi se informacija lokalnega značaja izgubila. Poleg tega je glede na dostopne meritve v delovnem območju v okolici ravnotežne krivulje lažje identificirati lokalne modele in jih dopolniti z redkimi meritvami s področja stran od ravnotežja.

Kot negativno lastnost metode lahko navedemo veliko računsko zahtevnost pri optimizaciji hiperparametrov, ki raste s tretjo potenco velikosti učne množice. Naslednja slabost, ki jo je pokazalo naše vrednotenje je, da razširjanje variance skozi GP model z vključenimi lokalnimi modeli višjega reda še ni dovolj raziskano.

Z višanjem reda dinamičnega modela oziroma s povečanjem števila vhodnih regresorjev se zelo poveča ne samo obseg modela, ampak tudi kompleksnost izbire podatkov za identifikacijo.

Modeliranje dinamičnih sistemov za Gaussovimi procesi je relativno novo področje, vendar so nekatere prednosti pristopa očitne. Na njihovi podlagi se obe-tajo uporabe teh modelov za vodenje dinamičnih sistemov, za diagnostiko napak in povsod drugod, kjer potrebujemo modele dinamičnih sistemov.

Poleg preizkusa potencialnih uporab, menimo, da je treba delo nadaljevati tudi na področju izdelave programske opreme in izboljševanju učinkovitosti algoritmov, da bi lahko opravljali praktično učinkovite postopke identifikacije dinamičnih sistemov z Gaussovimi procesi. Zavedamo se, da smo s predstavljenim vrednotenjem naredili šele majhen korak v razvoju te identifikacijske metode, vendar upamo, da bodo na novo odprta vprašanja pripomogla k učinkovitejšemu

nadaljnemu delu.

Literatura

- [1] G. Dolanc, Đ. Juričić, A. Rakar, J. Petrovčič, D. Vrančič (1997). Three-tank benchmark test, Copernicus Project, report CT94-02337, Inštitut Jožef Stefan, Ljubljana.
- [2] M.N. Gibbs (1997). Bayesian Gaussian Processes for Regression and Classification, Ph.D. Thesis, Cambridge University, Cambridge.
- [3] A. Girard, C.E. Rasmussen, R. Murray-Smith (2002). Gaussian Process priors with Uncertain Inputs: Multiple-Step-Ahead Prediction, Technical report, DCS TR-2002-119, University of Glasgow, Glasgow.
- [4] A. Girard, C.E. Rasmussen, J. Quinero Candela, R. Murray-Smith (2002). Multiple-step ahead prediction for non linear dynamic systems - A Gaussian Process treatment with propagation of the uncertainty, v Advances in Neural Information Processing Systems, urd. S. Becker, S. Thrun, K. Obermayer. Volume 15, MIT Press, 545-552.
- [5] A. Girard (2004). Approximate Methods for Propagation of Uncertainty with Gaussian Process Models, Ph.D. Thesis, University of Glasgow, Glasgow.
- [6] G. Gregorčič, G. Lightbody (2002). Gaussian Processes for Modelling of Dynamic Non-linear Systems, Irish Signals and Systems Conference, Cork.
- [7] G. Gregorčič, G. Lightbody (2003). An affine local Gaussian Process model network, Proceedings 16th International Conference on Systems Engineering ICSE 2003, Coventry University, 206-212.
- [8] J. Kocijan, A. Girard, D.J. Leith (2003). Incorporating Linear Models in Gaussian Process Model, Report DP-8895, IJS, Ljubljana.
- [9] J. Kocijan, B. Banko, B. Likar, A. Girard, R. Murray-Smith, C.E. Rasmussen (2003). A case based comparison of identification with neural networks and Ga-

- ussian process models, v International Conference on Intelligent Control Systems and Signal Processing, Faro, 137-142.
- [10] J. Kocijan, A. Girard, B. Banko, R. Murray-Smith (2003) Dynamic Systems Identification with Gaussian Processes, Proceedings 4th Mathmod Dunaj, sledi objava v Mathematical and Computer Modelling of Dynamic Systems.
- [11] D.J.Leith, W.E.Leithead, E. Solak, R. Murray Smith (2002). Divide and conquer identification: Using Gaussian process to combine derivative and non-derivative observations in a consistent manner, v Conference on Decision and Control, Las Vegas.
- [12] D.J.C. MacKay (1998). Introduction to Gaussian Processes, v Neural Networks and Machine Learning, Urd. C.M. Bishop, NATO ASI Series, 133-166.
- [13] D.J.C. MacKay (2003). Information Theory, Inference and Learning Algorithms, Cambridge University Press, Cambridge.
- [14] D. Matko (1998). Identifikacije, Založba FE in FRI, Ljubljana.
- [15] D. Matko, I. Škrjanc, G. Mušič (2000). Robustness of fuzzy control and its application to a thermal plant, v Mathematics and Computers in Simulation, 51, 245-255.
- [16] R. Murray-Smith, T.A. Johansen, R. Shorten (1999). On transient dynamics, off-equilibrium behaviour and identification in blended multiple model structures, European Control Conference ECC 1999, Karlsruhe, BA-14.
- [17] K.S. Narendra, K. Parthasarathy (1990). Identification and Control of Dynamical Systems Using Neural Networks, v IEEE Transactions on NN, Vol.1 No. 1, 4-27.
- [18] R.M. Neal (1995). Bayesian Learning for Neural Networks, Ph.D. Thesis, University of Toronto, Toronto.
- [19] A. O'Hagan (1978). On curve fitting and optimal design for regression (with discussion), v Journal of the Royal Statistical Society B, 40, 1-42
- [20] A. Rakar (2000). Odkrivanje napak v tehničnih sistemih z metodami aproksimativnega sklepanja, doktorska disertacija, Fakulteta za elektrotehniko, Univerza v Ljubljani, Ljubljana.
- [21] C.E. Rasmussen (1996). Evaluation of Gaussian Processes and Other Methods for NonLinear Regresion, PhD thesis, University of Toronto, Toronto.

- [22] E. Solak et al. (2003) Derivative observations in Gaussian Process Models of Dynamic Systems, v Neural Information Processing Systems NIPS 15, MIT Press, 529-536.

Izjavljam, da sem magistrsko nalogo izdelal samostojno pod vodstvom mentorja prof. dr. Juša Kocijana. Izkazano pomoč ostalih sem v celoti navedel v zahvali.

Ljubljana, september 2004