

Univerza v Ljubljani  
Fakulteta za elektrotehniko

Kristjan Ažman

# **Identifikacija dinamičnih sistemov z Gaussovimi procesi**

Doktorska disertacija

Mentor: prof. dr. Juš Kocijan

Ljubljana, 2007



”Izkušnost je tista čudovita stvar, ki ti omogoča prepoznati  
napako, ko jo storiš ponovno.”  
(Experience is that marvelous thing that enables you to recognize  
a mistake when you make it again.)

Franklin P. Jones



# Zahvala

*Iskreno se zahvaljujem mentorju prof. dr. Jušu Kocijanu za predstavitev zanimivega raziskovalnega področja in za vse predloge in nasvete, ki so mi pomagali tako na ožjem kot širšem področju raziskovalnega dela.*

*Iskreno se zahvaljujem tudi vodji odseka prof. dr. Stanku Strmčniku, ki mi je omogočil delovanje na odseku za Sisteme in vodenje Instituta Jožef Stefan, za ustvarjanje prijetnega vzdušja na odseku in za tehtne pripombe, ki so pomagale pri nastanku tega dela.*

*Prav tako gre zahvala vsem ostalim sodelavcem tega odseka, ki so s svojim znanjem in izkušnjami veliko prispevali k mojemu znanju in s pozitivnim odnosom poskrbeli za dobro počutje na delovnem mestu.*

*Zahvala gre tudi prof. dr. Dragu Matku, ki mi je že z diplomsko nalogo odprl vrata v zanimiv svet, ki je kasneje botroval nastanku tega dela.*

*Rad bi se zahvalil tudi prof. dr. Roderick Murray-Smithu za povabilo na Fakulteto za računalništvo Univerze v Glasgowu, kjer sem razširil znanje ne samo na tehničnih, ampak tudi na kulturnih področjih. Posebna zahvala gre tudi vsem članom laboratorija in njihovim prijateljem, ki so naredili moje bivanje v Glasgowu prijetno in zabavno.*

*Za financiranje mojega raziskovanja bi se rad zahvalil Agenciji za raziskovalno dejavnost Republike Slovenije in javnemu skladu Ad-Futura.*

*Zahvala gre tudi vsem prijateljem, ki so mi z ustvarjanjem prijetnega ozračja nesebično pomagali pri premagovanju spečnosti in mi s povzročanjem majhnih težav delali življenje bolj dinamično.*

*Posebna zahvala gre seveda družini, ki mi je vedno stala ob strani.*



# Povzetek

V delu je prikazana uporaba modelov na osnovi Gaussovih procesov, krajše GP modelov, za identifikacijo dinamičnih sistemov. GP model se kot orodje za identifikacijo ponuja predvsem za modeliranje nelinearnih sistemov.

GP model je verjetnostni, neparametrični model. Njegova poglobljena posebnost je, da je njegov izhod dan v obliki normalne (Gaussove) porazdelitve. Njeno srednjo vrednost lahko interpretiramo kot najbolj verjetno vrednost izhoda, varianca pa je mera zaupanja v to napoved.

Prvotno se je GP model uporabljal za modeliranje statičnih sistemov, medtem ko se za opisovanje dinamičnih sistemov uporablja razmeroma malo časa, zaradi česar je tudi manj literature o tovrstni uporabi. Ena izmed nalog disertacije je bila na enem mestu zbrati informacije o uporabi GP modela za identifikacijo dinamičnih sistemov. Naprej je postavljen okvir postopka identifikacije, nato so nekatere posamezne faze postopka predstavljene podrobneje: postavitve GP modela, vrednotenje modela in njegova simulacija.

Drugi sklop disertacije opisuje še eno dobro lastnost GP modela, tj. možnost vključevanja različnih vrst predznanja. Povzeti so znani rezultati o vključevanju linearnih lokalnih modelov in statične karakteristike, možnosti za vključevanje pa so razširjene z dognanji o vključevanju znanja o histerezi, o šumnem procesu na izhodu sistema in vključevanju znanja o strukturi procesa.

V tretjem sklopu disertacije je predstavljenih nekaj primerov za lažji prenos GP modela v praktično uporabo pri identifikaciji dinamičnih sistemov. Predstavljena je uporaba GP modela za filtriranje, napovedovanje odziva, načrtovanje vodenja in zaznavanje napak.

Rezultati kažejo, da je GP model lahko uporabno orodje za identifikacijo nelinearnih dinamičnih sistemov. Dobre lastnosti GP modela, ki so se pokazale med njegovo uporabo, so predvsem:

- varianca napovedanega izhoda sistema, ki je lahko uporabna pri načrtovanju vodenja, napovedovanju odziva, zaznavanju napak in pri vrednotenju modela;

- vključevanje različnih vrst predznanja v GP model, s čimer izboljšamo lastnosti GP modela in ga naredimo bolj uporabnega za inženirsko prakso in
- enostavnost uporabe.

Enostavnost uporabe GP modela je prav gotovo ena od lastnosti, ki bi lahko povečala njegovo uporabo v inženirski praksi. Za še večjo uporabnost bi poskrbel razvoj uporabniku prijaznejšega vmesnika ter razvoj in uporaba naprednejših in hitrejših algoritmov predvsem za učenje GP modela.

**KLJUČNE BESEDE:**

*nelinearni dinamični sistemi, identifikacija, Gaussovi procesi, model na osnovi Gaussovih procesov*

# Abstract

This thesis describes the use of a model based on Gaussian processes (the GP model) for dynamic system identification. As an identification tool the GP model is mainly for the identification of nonlinear dynamic systems.

The GP model is a probabilistic, nonparametrical model. Its main peculiarity is that the prediction is not given as a value but rather as a normal probability distribution, expressed in terms of the mean and variance. The mean value represents the most likely output and the variance can be viewed as a measure of its confidence.

The GP model was first used for the modelling of static systems, and only recently has it been proposed for modelling dynamic systems as well, which is the reason for the lack of literature on the topic. For this reason one of the objectives of the thesis was to gather information on the use of the GP model for dynamic system identification and present it in a consistent manner. In the first part of the thesis the general GP model identification procedure is presented, followed by a detailed presentation of some of the identification steps, namely, GP model selection, validation and simulation.

Another potentially very useful attribute of the GP model is explored in the thesis, *i.e.* the ability to incorporate various kinds of prior knowledge within the model. The incorporation of linear local models and static characteristic are already known and are presented in the thesis for completeness. To the list of possibilities for prior knowledge incorporation we added the knowledge about hysteresis, known noise process at the output of the system and the knowledge about the structure of the modelled system.

In the third part of the thesis, to enable a quicker and easier transfer of the GP model to practical use in dynamic system identification, some practical examples are presented, showing how to use the GP model for filtering, prediction, control design and fault detection.

The results of the work show that the GP model could be used as a tool for dynamic system identification. Several positive attributes of the GP model have been identified:

- the predicted variance of the system's output prediction can be used for control design, output prediction, fault detection and also for validating the model;
- the possibility to incorporate various kinds of prior knowledge into the GP model improves the model's attributes and makes it more attractive for engineering use;
- the model is simple to use.

Being able to use the GP model in a simple fashion is definitely one of the factors that could increase its use in the engineering community. This process could be speeded up with the design of a more easy-to-use user interface. Also, the employment of more advanced and quicker algorithms for the GP model training should be used, as this is the most time-consuming part of the identification process.

**KEY WORDS:**

*nonlinear dynamic systems, identification, Gaussian processes, Gaussian processes model*

# Originalni prispevki disertacije

- **Predlog metodologije za identifikacijo na osnovi Gaussovih procesov in njena uporaba na praktičnem primeru.**

Modeli na osnovi Gaussovih procesov so se najprej uporabljali za modeliranje statičnih sistemov in so se šele v zadnjem času začeli uporabljati za modeliranje dinamičnih sistemov. Te delne rezultate in prikaze modeliranja dinamičnih sistemov smo povezali v celosten postopek za identifikacijo dinamičnih sistemov z modeli na osnovi Gaussovih procesov in njihovimi inačicami z različnimi oblikami vključenega predznanja. Postopek identifikacije smo predstavili z običajnimi fazami in podrobneje opredelili in opisali tiste faze, ki so specifične pri uporabi modela na osnovi Gaussovih procesov: postavitve modela, vrednotenje in simulacija modela. Pri postavitvi modela smo podali smernice za izbiro kovariančne funkcije in regresorjev oziroma reda modela. Pri vrednotenju in simulaciji smo podali kriterijske funkcije, ki jih uporabljamo kot kvantitativno merilo, in opisali, kako uporabljamo vhodno/izhodni odziv modela kot kvalitativno merilo, s poudarkom na varianci in specifični simulacije. Našteli in pokazali smo prednosti in slabosti metode in svetovali kdaj uporabiti modele na osnovi Gaussovih procesov in kdaj njihove inačice. Identifikacijo dinamičnih sistemov smo prikazali na praktičnem primeru, pri čemer smo glavni poudarek namenili vrednotenju modela. Podali smo tudi predlog osnutka programske opreme za identifikacijo dinamičnih sistemov.

- **Vključevanje lokalnih modelov in predznanj v identifikacijski postopek.**

Vključevanje predznanja v model na osnovi Gaussovih procesov lahko pomembno vpliva na njegove lastnosti; model lahko zaradi vključenega predznanja bolje napoveduje obnašanje procesa ali pa se pohitri delovanje modela. Predstavili smo dva načina, s katerima splošno vključujemo predznanje: spreminjanje kovariančne funkcije in spreminjanje vhodno/izhodnih podatkov. Povzeli smo vključevanje lineranih lokalnih modelov in statične karakteristike ter ju za zaznavanje napak ovrednotili na praktičnem primeru. Poleg tega smo predlagali in ilustrirali, kako vključevati znanje o histerezi in znanje o šumnem procesu na izhodu sistema kot specifičnima oblikama predznanja. Posebej smo prikazali še, kako uporabimo znanje o strukturi opisovanega procesa.

- **Modeliranje in identifikacija dinamičnih procesov z modeli na osnovi Gaussovih procesov z vnaprej določeno strukturo.**

Znanje o strukturi opisovanega procesa lahko uporabimo za gradnjo parametričnega modela s spremenljivimi parametri, ki smo ga imenovali

model na osnovi Gaussovih procesov z vnaprej določeno strukturo. Prednost takega modela pred drugimi izpeljankami modelov na osnovi Gaussovih procesov je njegova parametričnost, ki se lahko izkaže kot uporabna lastnost npr. pri načrtovanju vodenja. Metodo smo uporabili za identifikacijo sistema dveh posod in procesa za pripravo plina. Identificirana modela smo uporabili za načrtovanje vodenja.

# Vsebina

<b>1</b>	<b>Uvod</b>	<b>1</b>
<b>2</b>	<b>Model na osnovi Gaussovih procesov</b>	<b>9</b>
2.1	Osnovni pojmi in delovanje . . . . .	9
2.2	Primer uporabe GP modela na statičnem regresijskem problemu .	16
<b>3</b>	<b>Identifikacija dinamičnih sistemov z GP modelom</b>	<b>19</b>
3.1	Uvod v identifikacijo nelinearnih dinamičnih sistemov . . . . .	19
3.1.1	Pristopi k identifikaciji nelinearnih dinamičnih sistemov . .	20
3.1.2	Problemi identifikacije nelinearnih dinamičnih sistemov . .	21
3.2	Identifikacija dinamičnih sistemov z GP modelom . . . . .	25
3.2.1	Lastnosti GP modela . . . . .	25
3.2.2	Predstavitev dinamičnih sistemov z GP modelom . . . . .	27
3.3	Postopek identifikacije . . . . .	28
3.3.1	Definicija namena modela . . . . .	29
3.3.2	Postavitev modela . . . . .	29
3.3.3	Načrtovanje eksperimenta . . . . .	29
3.3.4	Eksperiment in obdelava podatkov . . . . .	31
3.3.5	Učenje modela . . . . .	32
3.3.6	Vrednotenje modela . . . . .	32

3.4	Postavitev GP modela . . . . .	34
3.4.1	Izbira kovariančne funkcije . . . . .	34
3.4.2	Regresorji . . . . .	36
3.5	Simulacija GP modelov dinamičnih sistemov . . . . .	37
3.5.1	Simulacija brez propagacije negotovosti . . . . .	38
3.5.2	Simulacija s propagacijo negotovosti . . . . .	40
3.5.3	Numerična propagacija negotovosti z uporabo MCMC metod . . . . .	41
3.5.4	Smotrnost propagacije negotovosti . . . . .	44
3.6	Vrednotenje . . . . .	45
3.7	Primer identifikacije modela bioreaktorja . . . . .	47
3.8	Primer vrednotenja – Beneška laguna . . . . .	54
<b>4</b>	<b>Vključevanje predznanja v GP model</b>	<b>59</b>
4.1	Načini vključevanje predznanja . . . . .	60
4.1.1	Spreminjanje kovariančne funkcije/matrike . . . . .	60
4.1.2	Spreminjanje vhodno/izhodnih podatkov . . . . .	61
4.2	GP model z vključenimi lokalnimi modeli . . . . .	65
4.2.1	Linearni lokalni modeli . . . . .	66
4.2.2	Vključevanje lokalnih modelov v GP model . . . . .	67
4.2.3	Napovedovanje z uporabo LMGP modela in njegova simulacija . . . . .	72
4.2.4	Vrednotenje . . . . .	73
4.2.5	Primer identifikacije dinamičnega sistema drugega reda z LMGP modelom . . . . .	73
4.3	GP model z vnaprej določeno strukturo . . . . .	81
4.3.1	Modeliranje . . . . .	82

4.3.2	Simulacija FSGP modela . . . . .	85
4.3.3	Postopek identifikacije FSGP modela . . . . .	87
4.3.4	Primer identifikacije dinamičnega sistema drugega reda s FSGP modelom . . . . .	90
4.4	Ilustracija drugih načinov vključevanja predznanja v GP model . . . . .	96
4.4.1	Vključevanje predznanja o statični karakteristiki . . . . .	97
4.4.2	Vključevanje znanja o histerezi v GP model . . . . .	97
4.4.3	Poznavanje šumnega procesa na izhodu – vključevanje barvnega šuma v kovariančno matriko . . . . .	102
<b>5</b>	<b>Praktični primeri uporabe GP modelov</b>	<b>107</b>
5.1	Uvod . . . . .	107
5.2	Glajenje odzivov z GP modelom . . . . .	109
5.3	Identifikacija biološke čistilne naprave za napovedovanje odziva . . . . .	113
5.4	Identifikacija sistema dveh posod s FSGP modelom za načrtovanje vodenja . . . . .	123
5.5	Identifikacija procesa priprave plina s FSGP modelom za načrtovanje vodenja . . . . .	135
5.6	Identifikacija sistema dveh posod za zaznavanje napak . . . . .	145
5.7	Kdaj uporabiti GP model? . . . . .	158
<b>6</b>	<b>Zaključek</b>	<b>161</b>
	<b>Literatura</b>	<b>165</b>
	<b>Dodatek A - Nekateri uporabljeni matematični izrazi</b>	<b>177</b>
A.1	Gaussova porazdelitev in nekatere lastnosti . . . . .	177
A.2	Produkt dveh Gaussovih distribucij . . . . .	178
A.3	Izrek o invertiranju matrik . . . . .	178

A.4	Statistični momenti . . . . .	179
A.5	Statistični momenti mešanice Gaussovih porazdelitev . . . . .	180
A.6	Toeplitzova matrika . . . . .	181
<b>Dodatek B - Nekatere izpeljave</b>		<b>183</b>
B.1	Simulacija GP modela z analitično propagacijo negotovosti . . . . .	183
B.2	Simulacija FSGP modela . . . . .	190
<b>Dodatek C - Algoritmi za delo z GP modeli</b>		<b>193</b>
C.1	Simulacija GP modela z numerično propagacijo negotovosti . . . . .	193
C.2	Konstrukcija šumnega dela kovariančne matrike . . . . .	196
C.3	Simulacija FSGP modela . . . . .	198
C.4	Napovedovanje parametrov FSGP modela . . . . .	199
<b>Dodatek D - Interpretacija Takensovega teorema</b>		<b>201</b>
<b>Dodatek E - Opisi procesov</b>		<b>203</b>
E.1	Hidravlična modelna naprava . . . . .	203
E.2	Proces priprave plina . . . . .	207
<b>Dodatek F - Zasnova programskega modula za identifikacijo z GP modeli</b>		<b>213</b>

# Poglavje 1

## Uvod

Modele uporabljamo za opis sistemov; kot *sistem* razumemo omejeno ureditev procesov, ki vplivajo drug na drugega, pri čemer *proces* označuje pretvorbo in/ali prenos snovi, energije in/ali informacije. Osnovna lastnost modelov je, da poskušajo združiti opazovanja sistema v neki vzorec, ki ima enake lastnosti kot opazovani sistem. Poznamo več vrst modelov, vendar nas za opisovanje dinamičnih sistemov na področju systemske teorije zanimajo predvsem matematični modeli. [69]

Matematične modele lahko uporabljamo za napovedovanje obnašanja sistema, načrtovanje vodenja, zaznavanje napak itd. Kadar modelov sistemov ni mogoče dobiti z uporabo teoretičnega modeliranja (angl. *white-box modelling*) na osnovi fizikalnih, kemijskih ali drugih principov, na katerih temelji obnašanje sistema, si moramo pomagati z eksperimentalnim modeliranjem ali identifikacijo. Pri tem navadno predpostavimo neko parametrično strukturo modela, parametre pa določimo iz podatkov z uporabo izbranega kriterija. Primerno strukturo modela določimo iz predznanja o sistemu. Če poznamo strukturo sistema, določimo modelu enako strukturo in optimiramo njegove parametre. Če je znanje o strukturi sistema omejeno, poskusimo izbrati strukturo modela tako, da ta odraža znane značilnosti sistema in je model zmožen sistem do neke (zadovoljive) mere opisati. V obeh primerih zaradi kombinacije teoretičnega in eksperimentalnega modeliranja govorimo o kombiniranem modeliranju ali o identifikaciji s predznanjem o sistemu (angl. *grey-box modelling*). Seveda obstaja tudi možnost, da je znanje o sistemu pred identifikacijo zelo pomanjkljivo ali ga sploh ni. V tem primeru moramo strukturo modela izbrati dovolj fleksibilno, da je ta zmožna opisati raznovrstne potencialne strukture sistema: govorimo o eksperimentalnem modeliranju ali o identifikaciji brez predznanja oz. identifikaciji z modelom črne škatle (angl. *black-box modelling*).

V inženirski praksi se pri identifikaciji dinamičnih sistemov dostikrat predpostavlja, da je identificirani sistem linearen. Ta predpostavka sicer v praksi

nikoli ne drži, vendar se velikokrat izkaže kot dovolj dobra aproksimacija. Teorija identifikacije linearnih dinamičnih sistemov je uveljavljena in dobro znana. Če nasprotno upoštevamo nelinearnost sistema, večine teoretičnih rezultatov, znanih iz identifikacije linearnih dinamičnih sistemov, ne moremo uporabiti. S tem postane problem identifikacije sistema precej težji in poseči moramo po modelih, ki so zmožni oponašati sisteme nelinearnih struktur. Primeri takih modelov so umetne nevronske mreže, mehki modeli, mreža lokalnih modelov, valčki, metoda podpornih vektorjev ipd. Eden izmed razlogov za težjo identifikacijo nelinearnih sistemov je poleg manjšega nabora metod tudi izbira primerne strukture modela, saj se število potencialno uporabnih struktur nelinearnih modelov močno poveča. Preprosto povedano: neki sistem je lahko nelinearen na več načinov, kar navadno poveča kompleksnost modela, s katerim želimo sistem opisati. To pri opisovanju večjih in/ali kompleksnejših sistemov pripelje do problemov, kot so netransparentnost (tudi nepreglednost), “prekletstvo dimenzije” itd. Nekaterim problemom se z uporabo določenih pristopov sicer lahko ognemo, vendar pri tem navadno naletimo na druge probleme. To je eden izmed vzrokov, zakaj identifikacija nelinearnih dinamičnih sistemov še zdaleč ni do konca dognana in so raziskovalci na tem področju še zelo aktivni.

V disertaciji nas bo zanimala *identifikacija nelinearnih dinamičnih sistemov z modelom na osnovi Gaussovih procesov*, ki se je kot ena izmed potencialnih metod za identifikacijo nelinearnih dinamičnih sistemov pojavila konec devetdesetih let. Ta razmeroma nova metoda daje verjetnostni, neparometrični model neznanega sistema. Model na osnovi Gaussovih procesov ali krajše GP model je neparometričen – neznanega sistema ne poskuša opisati s prilagajanjem parametrov (navadno velikega števila) baznih funkcij, ki sestavljajo model, kot je to značilno npr. za umetne nevronske mreže. Sestavljen je iz vhodno-izhodnih podatkov, ki opisujejo obnašanje opisovanega sistema in jih model uporablja za napovedovanje, in kovariančne funkcije, ki pove, v kakšni medsebojni odvisnosti so ti podatki oz. kakšne funkcije so verjetneje uporabljene pri opisu sistema. Izhod iz modela je verjetnostna porazdelitev v obliki Gaussove porazdelitve, pri čemer je srednja vrednost najbolj verjetna vrednost izhoda, varianco pa lahko intepretiramo kot zaupanje v to napoved. Izražanje zaupanja v napoved je lastnost, po kateri se model na osnovi Gaussovih procesov najbolj loči od večine drugih metod za identifikacijo dinamičnih sistemov.

GP model je že dolgo znan na področju geostatistike, kjer je po Krige-ju [52] poznan pod imenom “kriging”. Kot orodje za reševanje regresijskih problemov ga je leta 1978 predstavil O’Hagan [81], popularnost v krogu ljudi, ki se ukvarjajo s strojnim učenjem, pa je v devetdesetih letih prejšnjega stoletja pridobil najprej z deli Neila [79], ki je pokazal, kakšna je povezava med GP modelom in umetnimi nevronskimi mrežami, in Rasmussena [90], ki je model umestil v Bayesov okvir,

---

ter nato Gibbsa [26] in Williamsa [120]. Več o razvoju GP modela najdemo npr. v [65].

GP model se lahko uporablja za reševanje problemov klasifikacije, v katerih je izhod modela neka vrednost iz končne (ponavadi majhne) množice, in za reševanje problemov regresije, v katerih je izhod modela zvezen. V disertaciji se bomo omejili izključno na uporabo GP modela za reševanje regresijskih problemov, med katere spada tudi identifikacija dinamičnih sistemov.

Ker je GP model soroden umetnim nevronske mrežam, a ima njegova uporaba pred uporabo umetnih nevronske mrež nekatere prednosti, se je začelo razmišljati o uporabi GP modelov (tudi) na področjih, kjer so bile umetne nevronske mreže že uspešno uporabljene. Uporabo oz. raziskovanje uporabe GP modela v identifikaciji dinamičnih sistemov je spodbudil evropski projekt MAC (2000-2004), katerega partner je bil z Odsekom za sisteme in vodenje tudi Institut Jožef Stefan. Rezultati tega projekta so bila prva objavljena dela s področja identifikacije dinamičnih sistemov z GP modelom. Na simulacijskih primerih je bila predstavljena identifikacija dinamičnih sistemov in simulacija dobljenih modelov brez propagacije negotovosti [33] in z njo [27, 28, 47]. Prav tako je bila na simulacijskih primerih prikazana primerjava med GP modelom in umetno nevronske mrežo [45] ter med GP modelom in mrežo lokalnih modelov [32]. GP modeli so bili uporabljeni tudi kot gradniki mreže lokalnih modelov [34]. Uporaba GP modela za vodenje z uporabo napovedane variance izhoda je bila na ilustrativnih primerih predstavljena v [50, 76, 77, 95].

Vključevanje predznanja v neki model je pomembno s stališča uporabnika-inženirja, saj dodatna informacija lahko pripomore k boljšemu razumevanju modela. GP model s svojo zgradbo ponuja možnosti za vključevanje predznanja različnih vrst in oblik. V projektu MAC je bilo največ raziskav opravljenih v zvezi z vključevanjem linearnih lokalnih modelov v GP model, npr. [46, 48, 49, 99], proučevano pa je bilo tudi vključevanje znanja o barvnem šumu na izhodu statičnega sistema [73]. GP model z vključenimi lokalnimi modeli je bil uporabljen tudi za načrtovanje vodenja [49].

Spodbudni rezultati teh raziskav na področju identifikacije dinamičnih sistemov z GP modelom so povzročili nadaljevanje dela. Tako je bil GP model z vključenimi lokalnimi modeli uporabljen za identifikacijo iz meritev na modelni napravi [4, 7], na industrijski napravi pa je bila izvedena identifikacija GP modela in načrtovanje vodenja na njegovi podlagi [61, 62]. GP model je bil skupaj z umetno nevronske mrežo in modelom na osnovi mehke logike uporabljen tudi za napovedovanje koncentracije ozona v ozračju [31].

Rezultati opravljenih raziskav so nakazovali, da bi bil GP model lahko privlačen za inženirsko prakso zaradi več lastnosti, povzetih v navedenih ugotovitvah:

- Je enostaven za uporabo, saj se je v primeru identifikacije dinamičnih sistemov potrebno na podlagi znanja o procesu odločiti le za ustrezno kovariančno funkcijo in določiti regresorje modela.
- Deluje razmeroma dobro tudi z malo podatki, ki so na voljo za identifikacijo, kar je še posebno pomembno, če želimo opisati področja, za katera je tipično pomanjkanje podatkov, npr. zunajravnotežna področja delovanja dinamičnih sistemov.
- V model je mogoče vključevati znanje različnih vrst, npr. lokalne modele, statično karakteristiko, znanje o šumu itd.
- GP model deluje dobro tudi v primerih, ko so podatki, ki opisujejo sistem, pošumljeni.
- Napoved modela poleg napovedi izhoda sistema vključuje tudi zaupanje v to napoved. To informacijo lahko koristno uporabimo npr.:
  - v vodenju, pri katerem lahko regulator načrtujemo tako, da sistema ne pripelje v slabo opisana področja [13, 76],
  - v zaznavanju napak, pri katerem z danim zaupanjem v napoved dobimo dodatno informacijo [40] in
  - pri identifikaciji sistema, pri kateri si z zaupanjem v napoved modela pomagamo pri vrednotenju modela ali oceni kvalitete identifikacijskih podatkov.

## Namen disertacije

V disertaciji smo želeli zbrati znanje o identifikaciji dinamičnih sistemov z GP modeli in ga razširiti na področjih, kjer je bilo omejeno, a je pomembno za identifikacijo dinamičnih sistemov, z namenom, da povečamo dosegljivost metode za praktično uporabo. Obe področji, tako identifikacija nelinearnih dinamičnih sistemov kot uporaba GP modelov za reševanje regresijskih problemov, sta vsako posebej sicer že znani in opisani, vendar je identifikacija dinamičnih sistemov z GP modelom kot sinteza obeh področij še nepopolno raziskana. Doslej znane teoretične rezultate smo želeli zbrati in za potencialne uporabnike postaviti uporaben okvir za identifikacijo dinamičnih sistemov z GP modelom. Da bi bila metoda bolj dostopna, smo jo želeli prikazali in ovrednotili na nekaterih praktičnih primerih izbranih dinamičnih sistemov.

V disertaciji smo želeli tudi razdelati postopek identifikacije dinamičnih sistemov z GP modelom in ga predstavili po posameznih fazah; poudarjene bi bile nekatere uporabne lastnosti metode in nekatere lastnosti, ki delajo metodo identifikacije

dinamičnih sistemov z GP modelom drugačno v primerjavi z bolj uveljavljenimi modeli.

Prav tako smo v disertaciji želeli zbrati in razširiti znanje o različnih oblikah vključevanja predznanja zaradi njegove pomembnosti v praktičnih inženirskih aplikacijah. Želeli smo raziskati možnosti za vključevanje predznanja o histerezi in za vključevanje znanja o znanem barvnem šumu na izhodu dinamičnega sistema ter razširiti znanje o vključevanju linearnih lokalnih modelov s praktičnimi primeri. Zanimalo nas je tudi, kako izkoristiti predznanje o strukturi opisovanega sistema za razvoj parametričnega modela, ki bi temeljil na GP modelu, saj daje uporaba parametričnega modela na voljo širši nabor metod za načrtovanje vodenja.

## Zgradba

Po uvodu smo v drugem poglavju najprej predstavili osnovne pojme GP modela, njegovo delovanje in uporabo v regresijskih problemih. Potegnili smo osnovne primerjave z drugimi vrstami modelov in prikazali njegovo delovanje na primeru reševanja statičnega regresijskega problema.

V tretjem poglavju smo se posvetili identifikaciji dinamičnih sistemov z GP modelom. Na začetku smo predstavili identifikacijo nelinearnih dinamičnih sistemov in metode, ki se zanjo največkrat uporabljajo skupaj z nekaterimi problemi, ki se ob tem pojavljajo.

Nato smo pokazali, kako je GP model mogoče uporabiti za opis dinamičnih sistemov in predstavili postopek identifikacije nelinearnih dinamičnih sistemov. Tiste faze postopka identifikacije, v katerih se postopek bolj spremeni zaradi uporabe GP modela, smo si ogledali natančneje: postavitev, simulacija in vrednotenje GP modela. Dali smo napotke za izbiro kovariančne funkcije in regresorjev GP modela, ogledali smo si, kakšne možnosti imamo pri simulaciji GP modela in kako v simulaciji GP modela propagiramo varianco skozi dinamični model, ter predstavili možnosti, ki se ponujajo pri vrednotenju GP modela zaradi dodatne informacije, ki jo nosi zaupanje v napoved izhoda sistema. Napotke smo orisali z dvema primeroma, pri čemer je v prvem več poudarka na postopku identifikacije, v drugem pa na vrednotenju.

V četrtem poglavju je predstavljeno (eksplicitno) vključevanje različnih vrst predznanja v GP model. Tega lahko vključujemo na dva načina: s spreminjanjem kovariančne funkcije oz. kovariančne matrike in s spreminjanjem vhodnih podatkov modela. Predstavili smo vključevanje linearnih lokalnih modelov, statične karakteristike, znanja o histerezi in znanja o dinamiki šuma na izhodu identificiranega sistema. Tu smo si ogledali tudi, kako določimo model z linearno

strukturo s spremenljivimi parametri, pri čemer za napovedovanje teh parametrov uporabljamo GP modele, če predpostavimo strukturo identificiranega procesa. Vse predstavljene možnosti za vključevanje so ilustrirane s preprostimi primeri.

V petem poglavju smo dali napotke za praktično uporabo GP modela in prikazali uporabo GP modela za reševanje problemov iz prakse modeliranja. Najprej je prikazana uporaba GP modela za glajenje podatkov na rezultatih identifikacije frekvenčne karakteristike realnega procesa. Sledi identifikacija biološke čistilne naprave za napovedovanje odziva, pri čemer smo si podrobno ogledali postopek identifikacije dinamičnih sistemov z GP modelom, poudarek smo namenili njegovemu vrednotenju. Na primeru identifikacije procesa priprave plina in sistema dveh posod smo si ogledali, kako uporabimo znanje o strukturi identificiranega procesa za tvorjenje GP modela z vnaprej določeno strukturo. Pridobljena modela smo nato uporabili še za vodenje obeh procesov. Sledi predstavitev uporabe GP modela z vključenimi lokalnimi modeli za zaznavanje napak na senzorjih izhoda, ki je prav tako predstavljena na sistemu dveh posod. Na koncu smo strnili napotke za praktično uporabo GP modela.

V zaključku smo povzeli poglobitve rezultate disertacije, še enkrat opisali najpomembnejše prispevke in dali nekaj napotkov za nadaljnje delo.

## Glavni prispevki disertacije

Prvi izmed glavnih prispevkov disertacije je predlaganje metodologije eksperimentalnega modeliranja dinamičnih sistemov z modeli na osnovi Gaussovih procesov, ki doslej v tej celoviti obliki še ni bila narejena. Postopek identifikacije je skupaj s svojimi prednostmi in slabostmi predstavljen v tretjem poglavju, največji doprinos pa poleg predlaganega okvira identifikacije predstavlja opis vrednotenja modela, pri čemer lahko koristno uporabimo varianco napovedi GP modela. Ta del disertacije izhaja iz dela, objavljenega v:

- [10] K. Ažman, J. Kocijan, Gaussian process model validation: biotechnological case studies, I. Troch, F. Breiteneker, uredniki, *Proceedings of the 5th Vienna Symposium on Mathematical Modelling – MathMod*, Dunaj, 2006.
- [12] K. Ažman, J. Kocijan, Application of Gaussian processes for black-box modelling of biosystems, *ISA Transactions*, Vol. 46, No. 4, str. 443–457, oktober 2007.

Drugi večji prispevek razširja možnosti vključevanja predznanja v GP model in je predstavljen v četrtem poglavju. Poleg že znanega in povzetega vključevanja

lokalnih modelov in statične karakteristike, je predstavljeno še vključevanje predznanja o histerezi in znanja o šumnem procesu na izhodu modela, predhodno predstavljeno v:

- [8] K. Ažman, J. Kocijan, Identifikacija dinamičnega sistema s histerezo z modelom na osnovi Gaussovih procesov, B. Zajc, A. Trost, uredniki, *Zbornik štirinajste elektrotehniške in računalniške konference ERK*, Volume A, str. 253–256, Portorož, 2005.
- [5] K. Ažman, Incorporating prior knowledge into Gaussian process model, *Proceedings of 6th International PhD Workshop on Systems and Control – A Young Generation Viewpoint*, Volume A, str. 253–256, Izola, 2005.
- [11] K. Ažman, J. Kocijan, Identifikacija dinamičnega sistema z znanim modelom šuma z modelom na osnovi Gaussovih procesov, B. Zajc, A. Trost, uredniki, *Zbornik petnajste elektrotehniške in računalniške konference ERK*, Volume A, str. 289–292, Portorož, 2006.

Prav tako je v tem poglavju predlagana metoda modeliranja nelinearnih dinamičnih sistemov z modeli na osnovi Gaussovih procesov z vnaprej določeno linearno strukturo, ki je predstavljena v:

- [9] K. Ažman, J. Kocijan, An application of Gaussian process models for control design, *UKACC International Control Conference*, str. 1092–1123, Glasgow, 2006.

V tretjem in četrtem poglavju predstavljeni postopki so preizkušeni na praktičnih primerih, rezultati preizkusov so opisani v petem poglavju. Z identifikacijo biološke čistilne naprave za napovedovanje odziva je prikazana metodologija identifikacije dinamičnih sistemov z modeli na osnovi Gaussovih procesov s poudarkom na vrednotenju [10, 12]. Nato je za identifikacijo dveh naprav z namenom vodenja uporabljen predlagan postopek modeliranja z modelom na osnovi Gaussovih procesov z vnaprej določeno strukturo [9]. Prav tako je na realnem primeru prikazana uporaba modela na osnovi Gaussovih procesov z vključenimi lokalnimi modeli za zaznavanje napak, kjer je glavni doprinos uporaba podatkov, dobljenih z meritvami na realni napravi, uporaba modela na osnovi Gaussovih procesov z vključenimi lokalnimi modeli in uporaba numerične propagacije variance za bolj točno oceno mere zaupanja v napovedi modela.



## Poglavje 2

# Model na osnovi Gaussovih procesov

V tem poglavju je predstavljen model na osnovi Gaussovih procesov ali krajše GP model. V prvem delu so predstavljeni osnovni pojmi: kako deluje, kako ga učimo, kako ga uporabljamo za napovedovanje in kako lahko interpretiramo rezultate. Na koncu poglavja je na enostavnem primeru predstavljena uporaba GP modela za reševanje statičnega regresijskega problema.

### 2.1 Osnovni pojmi in delovanje

#### Gaussovi procesi in model na njihovi osnovi

GP model se sicer uporablja tako za regresijo kot za klasifikacijo, vendar se bomo v disertaciji omejili na njegovo uporabo v regresijskih problemih. Značilnosti modeliranja z modelom na osnovi Gaussovih procesov so opisane npr. v [26, 66, 120], podrobnejšo razlago pa najdemo v [91].

Model na osnovi Gaussovih procesov je neparametrični, verjetnostni model [91] za opis sistemov. Namesto za modeliranje običajne omejitve na neki razred (parametriziranih) funkcij s tem pristopom *a priori* dopuščamo opis neznanega sistema z neskončno množico različnih funkcij. Pri tem dopuščamo večjo verjetnost funkcij, za katere menimo, da se pri opisu sistema verjetneje pojavljajo, npr. gladke, stacionarne, periodične.

Gaussovi procesi so naključni procesi. Naključni proces je posplošitev naključne spremenljivke na neki od neodvisnih spremenljivk odvisen prostor. Če je vrednost naključne spremenljivke v vsaki točki tega prostora porazdeljena po Gaussovi (normalni) porazdelitvi, takemu procesu pravimo *Gaussov proces* (GP) [65]. Povedano drugače: naključni proces je Gaussov proces, če je za vsak vektor

neodvisnih spremenljivk  $\mathbf{x}$  vrednost funkcije  $f(\mathbf{x})$  porazdeljena po Gaussovi (normalni) porazdelitvi.

Vhod v GP model so posamezne vrednosti neodvisnih spremenljivk, zbrane v vhodnem vektorju  $\mathbf{x}$ , medtem ko je izhod iz GP modela verjetnostna porazdelitev izhodne vrednosti  $f(\mathbf{x})$  pri danem vhodnem vektorju.

Za poljubni nabor  $N$  vhodnih vektorjev  $\mathbf{x}_i$ ,  $i = 1, \dots, N$ , je GP določen z vektorjem srednjih vrednosti  $\mathbf{m} = [m_1(\mathbf{x}_1) \dots m_N(\mathbf{x}_n)]^T$  in kovariančno matriko  $\mathbf{K}$ ,

$$\mathbf{K} = \begin{bmatrix} K_{11} & \dots & K_{1N} \\ \vdots & \ddots & \vdots \\ K_{N1} & \dots & K_{NN} \end{bmatrix} \quad (2.1)$$

kjer:

$$m_i(\mathbf{x}_i) = \mathbb{E}[f(\mathbf{x}_i)] \quad (2.2)$$

in so elementi kovariančne matrike  $K_{ij}$ , običajno dobljeni z neko *kovariančno funkcijo*  $C(\mathbf{x}_i, \mathbf{x}_j)$ , določeni kot:

$$\text{cov}(f(\mathbf{x}_i), f(\mathbf{x}_j)) = \mathbb{E}[(f(\mathbf{x}_i) - m(\mathbf{x}_i)) (f(\mathbf{x}_j) - m(\mathbf{x}_j))], \quad (2.3)$$

kjer je  $\mathbb{E}[x] = \int_{-\infty}^{\infty} xp(x)dx$  matematično upanje (povprečna vrednost) naključne spremenljivke  $x$  z verjetnostno porazdelitvijo  $p(x)$  (Dodatek A.4).

Če je porazdelitev neke množice spremenljivk Gaussova, je Gaussova tudi porazdelitev katerekoli podmnožice elementov te množice, kar imenujemo zahteva po konsistenci (angl. *consistency requirement*). To pomembno lastnost za delovanje GP modela vedno dosežemo, če so elementi kovariančne matrike  $\mathbf{K}$  GP-ja dobljeni s kovariančno funkcijo [91].

## Kovariančna funkcija

Vrednost kovariančne funkcije  $C(\mathbf{x}_i, \mathbf{x}_j)$  izraža medsebojno odvisnost vrednosti izhodov  $f(\mathbf{x}_i)$  in  $f(\mathbf{x}_j)$  na podlagi vrednosti vhodnih vektorjev  $\mathbf{x}_i$  in  $\mathbf{x}_j$ . Kovariančna funkcija je lahko različnih oblik, potrebno je le, da za poljubni nabor  $N$  vhodnih vektorjev  $\mathbf{x}_i$ ,  $i = 1, \dots, N$ , tvori ne-negativno definitno kovariančno matriko  $\mathbf{K}$ . Kovariančne funkcije so lahko stacionarne, nestacionarne, periodične itd., podrobneje bomo o njih govorili v podpoglavju 3.4.1. Za zdaj naj povemo, da kovariančne funkcije, ki določa obliko neznane funkcije  $f(\mathbf{x})$ , navadno ne poznamo vnaprej, lahko pa iz znanja o splošnih lastnostih funkcije  $f(\mathbf{x})$  sklepamo na njeno obliko.

Največ, zlasti kadar o lastnostih funkcije  $f(\mathbf{x})$  ne vemo dovolj, se uporablja Gaussova kovariančna funkcija, ki izraža dve pogosti lastnosti procesov:

- gladkost, ki pove, da se bo izhod procesa z majhno spremembo vhoda razmeroma malo spremenil (medsebojna odvisnost dveh vrednosti izhodov je večja, če ustrezni vrednosti vhodov ležita blizu skupaj), in
- stacionarnost, pri kateri je kovarianca med dvema vhodnima vektorjema odvisna samo od njune medsebojne razdalje in ne tudi od njune absolutne lege v prostoru.

Pri njej je kovarianca med dvema izhodoma  $y_i = f(\mathbf{x}_i)$  in  $y_j = f(\mathbf{x}_j)$ :

$$K_{ij} = C(\mathbf{x}_i, \mathbf{x}_j) = v \exp \left[ -\frac{1}{2} \sum_{d=1}^D w_d (x_i^d - x_j^d)^2 \right]. \quad (2.4)$$

$D$  je dimenzija vhodnega prostora in določa dolžino vhodnega vektorja  $\mathbf{x}$ . Parametri  $v$  in  $w_d$ ,  $d = 1, \dots, D$ , so poljubno določljivi parametri kovariančne funkcije. Imenujemo jih *hiperparametri*<sup>1</sup> [79, 91]; s tem poudarimo, da so to parametri sicer neparametričnega modela<sup>2</sup>, ki določajo obliko neznane funkcije  $f(\mathbf{x})$ . Parameter  $v$  govori o varianci izhoda, parametri  $w_d$  pa odražajo pomembnost posamezne komponente vhodnega vektorja; večji je parameter  $w_d$ , vplivnejša je sprememba komponente vektorja  $x^d$  na vrednost izhoda. Da dana kovariančna funkcija tvori pozitivno definitno kovariančno matriko, morajo biti vsi parametri Gaussove kovariančne funkcije večji od nič.

## Modeliranje

Najlaže predstavimo delovanje GP modela na primeru. Vzemimo, da bi radi opisali neki sistem:

$$y = f(\mathbf{x}) + \epsilon, \quad (2.5)$$

kjer je  $\epsilon$  beli Gaussov šum z varianco  $v_0$ ,  $\epsilon \sim \mathcal{N}(0, v_0)$ . Na podlagi  $N$ -tih vhodno-izhodnih vzorcev, t.j. parov vodoravnih vektorjev  $\mathbf{x}_i$  in skalarjev  $y_i$ , zbranih v množici  $\mathcal{D} = \mathbf{X}, \mathbf{y}$ ,

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_N \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix},$$

<sup>1</sup>Neal [79] je pokazal, da je vnaprejšnja nevronska mreža z enim skritim nivojem z neskončnim številom nevronov in pri določenih porazdelitvah parametrov nevronske mreže enaka GP modelu. Hiperparametri določajo distribucijo vrednosti (sicer neskončnega števila) parametrov te nevronske mreže.

<sup>2</sup>Model je neparametričen, saj, kot bomo videli, za napovedovanje poleg hiperparametrov in kovariančne funkcije potrebujemo še informacijo o obnašanju sistema v obliki vhodno/izhodnih podatkov, uporabljenih pri modeliranju.

želimo določiti neznan vrednost izhoda  $y^*$  pri vrednostih vhodnega vektorja  $\mathbf{x}^*$ . V nadaljevanju v kontekstu GP modela  $N \times D$  matriko  $\mathbf{X}$  in  $N \times 1$  vektor  $\mathbf{y}$  označimo kot *učno množico*, saj jih uporabljamo za učenje (angl. *training*) GP modela. Posamezni vhodno/izhodni par  $(\mathbf{x}_i, y_i)$  iz te množice imenujemo tudi učni vektor oz. učna točka. Par  $(\mathbf{x}^*, y^*)$  označimo kot preizkusno oz. testno množico ali tudi kot *testni vhod/izhod*.

Učni izhodi  $y_i$ ,  $i = 1, \dots, N$ , predstavljajo vrednosti naključnih spremenljivk, izhajajočih iz Gaussovega procesa. Predpostavimo, da je izhod sistema gladek in da je sistem stacionaren ter za tvorjenje kovariančne matrike  $\mathbf{K}$  med posameznimi izhodi  $y_i$ ,  $i=1, \dots, N$ , uporabimo Gaussovo kovariančno funkcijo (2.4) z na začetku neznanimi parametri.

Dobimo:  $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_N)$  z elementi kovariančna matrike  $K_{N_{ij}} = \Sigma_{ij} + v_0 \delta_{ij}$ .  $\Sigma_{ij}$  so elementi kovariančne matrike, dobljeni s kovariančno funkcijo (2.4),  $v_0 \delta_{ij}$  pa opisuje vpliv šuma na izhodu procesa, kjer je  $\delta_{ij}$  Kroneckerjev operator ( $\delta_{ij} = 1$ ,  $i = j$ ;  $\delta_{ij} = 0$ ,  $i \neq j$ ). Ker smo predpostavili beli šum, so njegove vrednosti korelirane le same s seboj.

Z uporabo podatkov  $\mathcal{D} = \mathbf{X}, \mathbf{y}$ , ki jih imamo na voljo, bi radi določili neznan funkcijo  $f(\mathbf{x})$  iz enačbe (2.5). Funkcijo modeliramo z uporabo Bayesovega pristopa [66]:

$$p(f(\mathbf{x})|\mathbf{y}, \mathbf{X}) = \frac{p(\mathbf{y}|f(\mathbf{x}), \mathbf{X}) p(f(\mathbf{x}))}{p(\mathbf{y}|\mathbf{X})}. \quad (2.6)$$

Prvi izraz v števcu enačbe (2.6)  $p(\mathbf{y}|f(\mathbf{x}), \mathbf{X})$  predstavlja verjetnost učnih izhodov glede na funkcijo  $f(\mathbf{x})$  (in učne vhode  $\mathbf{X}$ ) in je v regresijskih problemih navadno predpostavljeno Gaussova [66]. Drugi izraz v števcu predstavlja a priori verjetnost posameznih funkcij, ki sestavljajo model. Ideja modela na osnovi Gaussovih procesov je, da funkcije  $f(\mathbf{x})$  ne parametriziramo, ampak določimo a priori verjetnosti direktno v funkcijskem prostoru [66]. Preprost primer take a priori verjetnosti je Gaussov proces.

Ker je neznan izhod  $y^*$  udejanjenje istega procesa kot učni izhodi  $\mathbf{y}$ , lahko zapišemo [26]:  $\mathbf{y}_{N+1} = \begin{bmatrix} \mathbf{y} \\ y^* \end{bmatrix} \sim \mathcal{N}(0, \mathbf{K}_{N+1})$ . Skupno kovariančno matriko  $\mathbf{K}_{N+1}$  vektorja  $\mathbf{y}_{N+1}$  lahko razdelimo:

$$\mathbf{K}_{N+1} = \begin{bmatrix} \begin{bmatrix} \mathbf{K} \end{bmatrix} & \begin{bmatrix} \mathbf{k}(\mathbf{x}^*) \end{bmatrix} \\ \begin{bmatrix} \mathbf{k}(\mathbf{x}^*)^T \end{bmatrix} & \begin{bmatrix} k(\mathbf{x}^*) \end{bmatrix} \end{bmatrix}. \quad (2.7)$$

Matrika  $\mathbf{K}$  je kovariančna matrika učnih podatkov,  $\mathbf{k}(\mathbf{x}^*)$  je vektor kovarianc med učnimi izhodi in testnim izhodom,  $k(\mathbf{x}^*)$  pa avtokovarianca testnega izhoda.

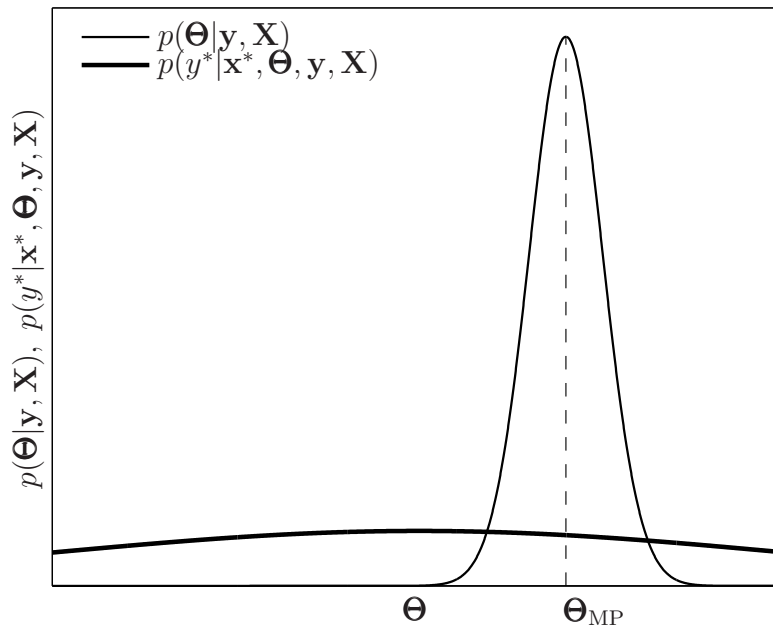
Verjetnostno porazdelitev  $p(y^*|\mathbf{x}^*, \mathbf{y}, \mathbf{X})$  vrednosti izhoda  $y^*$  lahko razdelimo na dva dela; na del, ki določa verjetnost učnih izhodov glede na učne vhode (angl. *marginal part*),  $p(\mathbf{y}|\mathbf{X}) \sim \mathcal{N}(\mathbf{0}, \mathbf{K})$ , in na pogojni del (angl. *conditional part*), ki glede na prvi del in vhod  $\mathbf{x}^*$  napoveduje verjetnostno porazdelitev izhoda  $y^*$ . Formalno zapisano je izračun porazdelitve izhodne verjetnosti odziva  $y^*$  [26, 65]:

$$p(y^*|\mathbf{x}^*, \mathbf{y}, \mathbf{X}) = \int p(y^*|\mathbf{x}^*, \boldsymbol{\Theta}, \mathbf{y}, \mathbf{X})p(\boldsymbol{\Theta}|\mathbf{y}, \mathbf{X})d\boldsymbol{\Theta}. \quad (2.8)$$

Običajno je ta integral analitično neizračunljiv, imamo pa na voljo dve alternativni [65]. Prva, pogotejša metoda je aproksimacija integrala z uporabo najbolj verjetnih vrednosti neznanih hiperparametrov  $\boldsymbol{\Theta}_{\text{MP}}$ :

$$p(y^*|\mathbf{x}^*, \mathbf{y}, \mathbf{X}) \approx p(y^*|\mathbf{x}^*, \boldsymbol{\Theta}_{\text{MP}}, \mathbf{y}, \mathbf{X}). \quad (2.9)$$

Uporabimo tiste vrednosti hiperparametrov  $\boldsymbol{\Theta}_{\text{MP}}$ , pri katerih je verjetnost učnih izhodov  $\mathbf{y}$  glede na vrednosti učnih vhodov  $\mathbf{X}$  in kovariančno funkcijo  $C(., .)$  največja. Pri tej metodi predpostavljamo, da ima porazdelitev izhodne verjetnosti glede na hiperparametre  $\boldsymbol{\Theta}$  oster vrh okoli vrednosti  $\boldsymbol{\Theta}_{\text{MP}}$  relativno glede na spremenljivost porazdelitve  $p(y^*|\mathbf{x}^*, \boldsymbol{\Theta}, \mathbf{y}, \mathbf{X})$  [26], slika 2.1.



**Slika 2.1:** Pri uporabi najbolj verjetnih parametrov  $\boldsymbol{\Theta}_{\text{MP}}$  predpostavljamo, da ima porazdelitev  $p(\boldsymbol{\Theta}|\mathbf{y}, \mathbf{X})$  relativno oster vrh glede na porazdelitev  $p(y^*|\mathbf{x}^*, \boldsymbol{\Theta}, \mathbf{y}, \mathbf{X})$ .

Najbolj verjetne vrednosti hiperparametrov  $\boldsymbol{\Theta}_{\text{MP}}$  dobimo z metodo največje podobnosti (angl. *maximum likelihood method*, ML). Da se izognemo optimizaciji

z omejitvami, za optimizacijo uporabimo logaritem porazdelitve učnih podatkov (angl. *log-marginal likelihood*):

$$\mathcal{L}(\boldsymbol{\Theta}) = \log(p(\mathbf{y}|\mathbf{X}, \boldsymbol{\Theta})) = -\frac{1}{2} \log(|\mathbf{K}|) - \frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} - \frac{N}{2} \log(2\pi), \quad (2.10)$$

kjer je  $\boldsymbol{\Theta} = [w_1 \dots w_D \ v \ v_0]^T$  vektor parametrov in  $\mathbf{K}$  kovariančna matrika učnih podatkov  $\mathcal{D}$ . Če je optimizacija izvedena z metodo konjugiranih gradientov (ali katero drugo gradientno metodo), je potreben še izračun odvodov po vseh hiperparametrih:

$$\frac{\partial \mathcal{L}(\boldsymbol{\Theta})}{\partial \Theta_i} = -\frac{1}{2} \text{trace} \left( \mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \Theta_i} \right) + \frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \Theta_i} \mathbf{K}^{-1} \mathbf{y}. \quad (2.11)$$

Ob vsakem koraku optimizacije je potrebno izračunati inverz kovariančne matrike  $\mathbf{K}^{-1}$ , kar je računsko zahtevno za velike  $N$ .

Druga alternativa za aproksimacijo integrala (2.8) je numerična integracija nad celotno porazdelitvijo hiperparametrov (MCMC metode [65]), dobljeno z optimizacijo verjetnosti učnih podatkov (2.10) [26]. Pri tem integracijo preko porazdelitve vektorja hiperparametrov  $p(\boldsymbol{\Theta}|\mathbf{y}, \mathbf{X})$  aproksimiramo s  $S$  vzorci porazdelitve  $\boldsymbol{\Theta}$  in integral (2.8) aproksimiramo s povprečno vrednostjo porazdelitev pri posameznih vzorčenih vrednostih hiperparametrov  $\boldsymbol{\Theta}_s$ :

$$p(y^*|\mathbf{x}^*, \mathbf{y}, \mathbf{X}) \approx \frac{1}{S} \sum_{s=1}^S p(y^*|\mathbf{x}^*, \boldsymbol{\Theta}_s, \mathbf{y}, \mathbf{X}). \quad (2.12)$$

Daljši opis numerične aproksimacije integrala (2.8) z numeričnimi metodami, problemov tega pristopa in njihovih rešitev je v [26] in tamkajšnjih referencah.

Obstaja še ena možnost za izbiro hiperparametrov modela, tj. navzkrižno vrednotenje (angl. *cross-validation*) [91]. Vrednosti hiperparametrov iščemo kot običajno, le da učne podatke razdelimo na  $k$  delov. Za učenje uporabimo  $k - 1$  delov, za vrednotenje pa tistega, ki ostane. Postopek ponovimo  $k$ -krat, vsakič z drugimi podatki za vrednotenje. Ekstremni primer je navzkrižno-vrednotenje z izpuščanjem enega podatka (angl. *leave-one-out cross-validation*, *LOO cross-validation*). Največji problem tega postopka je računska zahtevnost, saj moramo naučiti  $k$  modelov.

## Napovedovanje

Skupna porazdelitev  $p(\mathbf{y}_{N+1})$  je Gaussova; torej je Gaussova tudi pogojna porazdelitev  $p(y^*|\mathbf{x}^*, \mathbf{y}, \mathbf{X}) = \frac{p(\mathbf{y}_{N+1})}{p(\mathbf{y}|\mathbf{X})}$ . Po poenostavitvi [65, 120] kot napovedan izhod sistema (2.5) dobimo Gaussovo porazdelitev:

$$p(y^*|\mathbf{x}^*, \mathbf{X}, \mathbf{y}) = \mathcal{N}(\mu(\mathbf{x}^*), \sigma^2(\mathbf{x}^*)) \quad (2.13)$$

s srednjo vrednostjo  $\mu(\mathbf{x}^*)$  in varianco  $\sigma^2(\mathbf{x}^*)$ :

$$\mu(\mathbf{x}^*) = \mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1} \mathbf{y} \quad (2.14)$$

$$\sigma^2(\mathbf{x}^*) = k(\mathbf{x}^*) - \mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}^*), \quad (2.15)$$

kjer je  $\mathbf{k}(\mathbf{x}^*) = [C(\mathbf{x}_1, \mathbf{x}^*) \dots C(x_N, \mathbf{x}^*)]$  že omenjeni  $N \times 1$  kovariančni vektor med testnim izhodom in učnimi izhodi ter  $k(\mathbf{x}^*) = C(\mathbf{x}^*, \mathbf{x}^*)$  avtokovarianca testnega izhoda.

## Interpretacija

GP model je torej sestavljen iz dveh delov:

- iz parov vhodno/izhodnih učnih podatkov (točk)  $\mathcal{D}$ , ki predstavljajo obnašanje neznanega sistema, in
- kovariančne funkcije  $C(.,.)$  z znanimi oz. optimiranimi hiperparametri  $\Theta$ , ki pove, v kakšnem medsebojnem razmerju so podatki  $\mathcal{D}$ .

Ker GP model potrebuje informacijo o neznanu funkciji v obliki učnih vhodov in izhodov tudi po učenju, je model neparametričen. Hiperparametri namreč prek kovariančne funkcije samo povedo, kako se učna informacija uporabi za napovedovanje, ni pa v njih spravljena informacija o opisovani funkciji/sistemu kot v parametrih parametričnega modela.

Na vektor  $\mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1}$  v izrazu za srednjo vrednost napovedanega izhoda (2.14) lahko gledamo kot na vektor uteži, ki določa utežitev posameznih učnih izhodov  $y_i$  v  $\mathbf{y}$  glede na s hiperparametri  $w_d$  uteženo evklidsko razdaljo<sup>3</sup> med učnimi in testnim vhodnim vektorjem v vhodnem prostoru. Ta linearna kombinacija učnih izhodov (angl. *linear predictor*) se lahko razume kot glajenje v GP modelu vsebovane informacije o neznanem sistemu (učni podatki). Še drugače si lahko srednjo vrednost napovedi  $\mu(\mathbf{x}^*)$  predstavljamo kot linearno kombinacijo  $n$  jedrnih (angl. *kernel*) funkcij, usrediščenih v učnih točkah:  $\mu(\mathbf{x}^*) = \sum_{i=1}^N \alpha_i C(\mathbf{x}^*, \mathbf{x}_i)$  [91]. Izhod iz sistema je eden izmed vzorcev iz dobljene normalne porazdelitve (2.13).

Majhna varianca  $\sigma^2(\mathbf{x}^*)$  napovedane porazdelitve izhoda pomeni večje *zaupanje* v napoved in narobe. Če si ogledamo izraz za varianco, vidimo, da je sestavljen iz dveh delov [91]. Od prvega dela  $k(\mathbf{x}^*)$ , ki predstavlja a priori varianco GP, je odštet izraz  $\mathbf{k}(\mathbf{x}^*)^T \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}^*)$ . Ta zmanjša a priori varianco GP modela pri  $\mathbf{x}^*$  zaradi učnih podatkov in se večja z večjo kovarianco med učnimi in testnim

<sup>3</sup>tudi Mahalanobisova razdalja [84, 86].

vhodom. Preprosteje: “bližje” je testni vhod že znanim (učnim) vhodom v vhodnem prostoru, večje je zaupanje GP modela v točnost napovedi. Prav varianca, odvisna tudi od lege testnega vhoda glede na učne, je ena izmed glavnih prednosti GP modela pred drugimi modeli.

Delovanje GP modela si lahko predstavljamo tudi takole: denimo, da imamo naučen GP model z dimenzijo vhodnega prostora  $D$ . Učne točke, ki sestavljajo ta model, si lahko predstavljamo kot točke  $[\mathbf{x}_i, y_i]$  v  $D + 1$  dimenzionalnem prostoru. Te točke sestavljajo ogrodje, prek (oz. bolj natančno blizu, saj napoved GP modela pri učnem vhodu  $\mathbf{x}_i$  navadno ni povsem enaka vrednosti učnega izhoda  $y_i$ ) katerega so kot hiperpovršina “napete” srednje vrednosti napovedi GP modela za vse možne testne vhode  $\mathbf{x}^*$ , odgovarjajoče napovedi varianc oz. standardnih deviacij pa predstavljajo bolj ali manj majhen odmik od te hiperpovršine.

## 2.2 Primer uporabe GP modela na statičnem regresijskem problemu

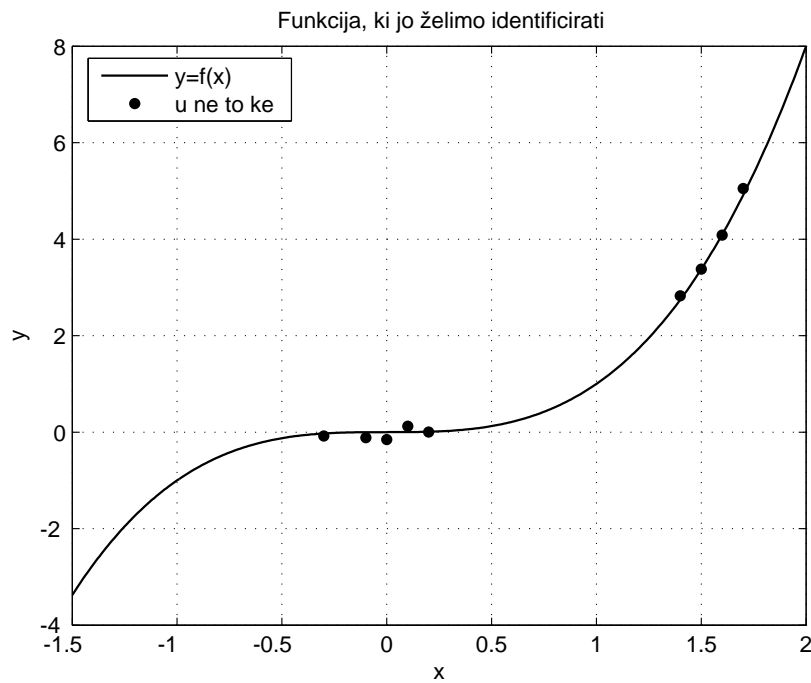
Ilustrirajmo uporabo GP modela na primeru. Želimo identificirati nelinearno funkcijo  $y = f(x)$ , odvisno od neodvisne spremenljivke  $x$ :

$$f(x) = x^3 + \epsilon \quad (2.16)$$

na intervalu  $x \in [-1.5, 2]$ . Varianca belega Gaussovega šuma  $\epsilon$  na izhodu je  $\sigma^2 = 0.01$ . Nelinearna funkcija je predstavljena z devetimi neenakomerno porazdeljenimi učnimi pari (točkami), generiranimi s funkcijo  $f(x)$ , ki predstavljajo vhodno/izhodno relacijo  $x/y$ . Funkcija  $y = f(x)$  in učne točke so prikazane na sliki 2.2.

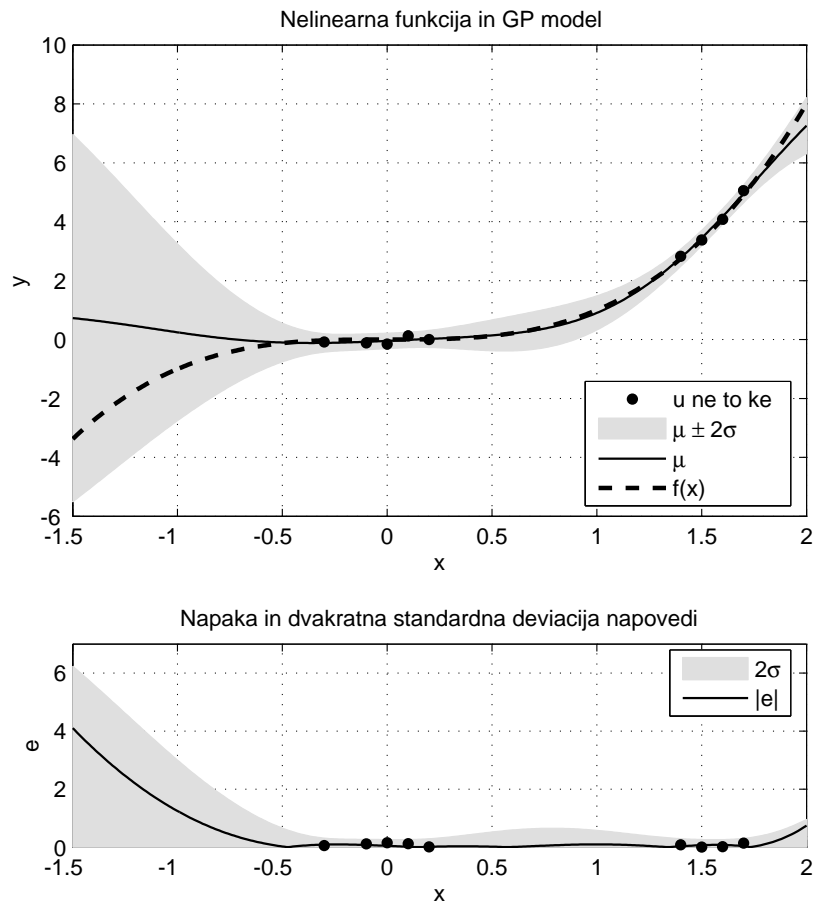
Za opis funkcije izberemo Gaussovo kovariančno funkcijo (2.4), dodamo ji del, ki modelira beli šum. Zaradi samo enega vhoda se kovariančna funkcija poenostavi v:  $C(x_i, x_j) = v \exp[-\frac{1}{2}w(x_i - x_j)^2] + v_0\delta_{ij}$ . Z učenjem hiperparametrov po metodi največje podobnosti, postopku, opisanem na str. 13, določimo tri hiperparametre:  $v = 26.4$ ,  $w = 0.96$  ter  $v_0 = 0.013$ .

Rezultati identifikacije so prikazani na sliki 2.3. Lahko opazimo, da model dobro aproksimira neznano funkcijo  $y = f(x)$  na področjih  $-0.3 < x < 0.2$  in  $1.4 < x < 1.7$ , ki so opisana z učnimi točkami. Prav tako model dobro napoveduje na področju  $0.2 < x < 1.4$ , ki leži med obema področjema. Čeprav je napoved dobra, model na tem področju prek relativno povečane variance napovedi izhoda izkazuje večjo negotovost v svojo napoved. Na območjih, kjer funkcija  $f(x)$  ni več opisana z učnimi podatki in mora GP model za napovedovanje ekstrapolirati, se napaka napovedi hitro veča, a se hkrati z njo veča tudi napovedana varianca GP



**Slika 2.2:** Nelinearna funkcija in učne točke za učenje GP modela

modela, kar je še posebej izrazito na področju  $x < -0.3$ . Zaradi manjšega šuma je manj opazno, kako GP model pošumljene učne vzorce z linearno kombinacijo zgladi za napovedovanje izhoda.



Slika 2.3: Primerjava nelinearne funkcije in opisa z GP modelom

## Poglavje 3

# Identifikacija dinamičnih sistemov z GP modelom

V tretjem poglavju je najprej predstavljena identifikacija nelinearnih dinamičnih sistemov in nato problemi, ki se pojavljajo pri njenem opisovanju z nekaterimi bolj pogostimi metodami. Potem so predstavljeni razlogi, zaradi katerih bi bilo smiselno za opis nelinearnih dinamičnih sistemov uporabiti GP model, predstavljen pa je tudi način, kako GP model uporabiti za opis dinamičnega sistema. Sledi predstavitev postopka identifikacije nelinearnih dinamičnih sistemov z GP modelom in podrobnejša predstavitev tistih faz identifikacijskega postopka, v katerih se identifikacija z GP modeli bolj loči od identifikacije z drugimi vrstami modelov: postavitve, simulacija in vrednotenje GP modela. Poglavje je zaključeno z dvema primeroma identifikacije nelinearnih dinamičnih sistemov z GP modelom, pri čemer je poudarek prvega primera na celotnem postopku identifikacije, medtem ko je poudarek drugega na vrednotenju identificiranega GP modela.

### 3.1 Uvod v identifikacijo nelinearnih dinamičnih sistemov

O modeliranju dinamičnih sistemov smo na splošno pisali že v uvodu disertacije. V tem poglavju nas bo zanimalo modeliranje dinamičnih sistemov, o katerih nimamo predznanja ali ga je zelo malo, torej bomo govorili o identifikaciji z modelom črne škatle. O vključevanju predznanja v GP model, ki GP model spremeni iz “črne” v “sivo škatlo”, bomo govorili v četrtem poglavju.

Kljub temu da so sistemi, ki jih želimo identificirati, navadno bolj ali manj nelinearni, lahko velikokrat s predpostavko o linearnosti sistema dovolj dobro identificiramo sistem, če le nelinearnost modeliranega sistema ni preveč izrazita. Identifikacija linearnih dinamičnih sistemov [64,69] je uveljavljena, že dolgo znana

in razmeroma enostavna. Če pa je nelinearnost identificiranega sistema prevelika, metode za identifikacijo linearnih sistemov navadno niso dovolj in moramo se zateči k metodam, s katerimi dobimo modele, zmožne opisa *nelinearnih dinamičnih sistemov*.

### 3.1.1 Pristopi k identifikaciji nelinearnih dinamičnih sistemov

Kljub temu da s predpostavko o nelinearnosti dinamičnega sistema omejimo krog potencialno uporabnih metod, je za njegovo identifikacijo na voljo še vedno razmeroma veliko pristopov. Mednje spadajo umetne nevronske mreže (angl. *artificial neural networks*, ANN) [23], modeli na osnovi mehke logike (angl. *fuzzy models*, tudi mehki modeli) [106], mreže lokalnih modelov (angl. *local model networks*, LMN) [74], metoda podpornih vektorjev (angl. *support vector machines*, SVM) [103], metoda relevantnih vektorjev (angl. *relevance vector machines*, RVM) [87], valčki (angl. *wavelets*) [104] itd. Primerjava različnih modelov črne škatle s poudarkom na strukturi modelov, možnih naborih regresorjev, vrstah preslikav itd. je predstavljena v [97], v spremljevalnem članku [39] pa so navedene potrebne matematične osnove. Pregled izbranih metod za modeliranje nelinearnih sistemov z modeli črne škatle je zbran npr. v [104], napotki za pomoč pri določanju strukture nelinearnega modela pa npr. v [85]. Metodam SVM, RVM, valčkom in tudi drugim (tu neomenjenim) metodam v disertaciji ne bomo namenjali pozornosti.

Najbolj množično se za identifikacijo nelinearnih dinamičnih sistemov uporabljajo umetne nevronske mreže in mehki modeli. Poznamo več vrst umetnih nevronskih mrež, npr. ANN z radialnimi baznimi funkcijami (RBF nevronske mreže, angl. *Radial Basis Functions ANN*) [86], vnaprejšnje (angl. *feed-forward*) ANN [86] itd. Poznamo tudi dva tipa mehkih modelov: Mamdani mehki model [67] in Takagi-Sugeno mehki model [106]. Prvi je po nekaterih lastnostih bolj podoben ANN, drugi pa mreži lokalnih modelov (LMN). Med posameznimi tipi modelov se odločamo glede na problem, ki ga želimo rešiti.

Predstavitev uporabnosti ANN in mehkih modelov za identifikacijo in vodenje, Narendra in Parthasarathy [78] za ANN in Mamdani ter Takagi in Sugeno [67,106] za mehke modele, je spodbudila razvoj teorije obeh področij [123,124], ta pa je razširjala uporabo obeh modelov v praksi. Tako sta danes obe vrsti modelov uporabljeni v širokem spektru praktičnih aplikacij, ki sega od procesne industrije [3,60], do uporabe v meteorologiji [24,31], ekonomiji [122], medicini [20] itd.

Obema metodama in omenjenim valčkom je skupen način opisa nelinearnega dinamičnega sistema. Metode namreč ponavadi uporabljajo veliko število baznih funkcij, ki zmorejo, združene v mrežo določene strukture, opisati poljubno nelinearnost. Po določitvi strukture, ki ni enostavno opravilo, se v postopku

identifikacije modelu določijo parametri vseh manjših sestavnih delov (baznih funkcij). Zakaj pride do problemov in zakaj razvijati alternativne metode, če so te strukture zmožne opisati poljubno nelinearnost?

### 3.1.2 Problemi identifikacije nelinearnih dinamičnih sistemov

Problemi identifikacije dinamičnih sistemov izhajajo iz dinamike sistemov, ki jih opisujemo, nelinearnosti sistema pa te probleme še povečajo in poudarijo. Zaradi nelinearnosti se namreč proces na različnih področjih delovanja obnaša različno in pojavi se potreba po bolj kompleksnih modelih. Če poznamo strukturo opisovanega sistema je problem nekoliko lažji, saj je potrebno iz podatkov samo določiti parametre neke že znane strukture modela. Če strukture sistema ne poznamo, se moramo zateči k uporabi metod modeliranja z modeli črne škatle, ki nas zanimajo v tem poglavju. Omejili smo se na dva tipična predstavnika metod za identifikacijo nelinearnih dinamičnih sistemov: na umetne nevronske mreže in na mrežo lokalnih modelov. Ker mreža lokalnih modelov po nekaterih lastnostih prekaša ANN in ker z njo sistem predstavimo podobno kot z izpeljankama GP modela, predstavljenima v četrtem poglavju, ji bomo namenili več pozornosti.

#### Umetne nevronske mreže

Kljub množični uporabi ANN imajo te poleg problema z določevanjem primerne strukture še dva večja problema [37]:

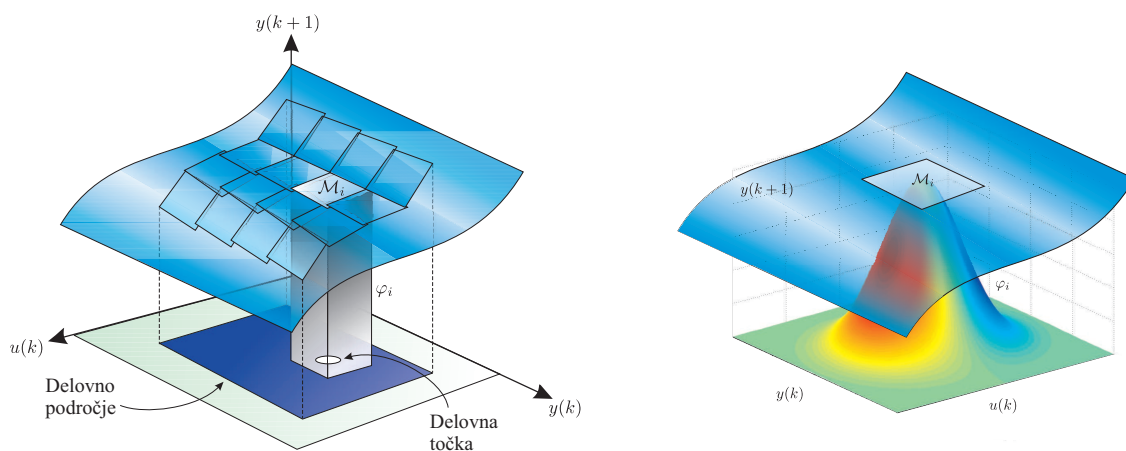
- netransparentnost (angl. *the lack of transparency*), struktura nevronske mreže ne odraža strukture neznanega sistema, in
- “prekletstvo dimenzije” (angl. *the curse of dimensionality*) [75], ki se izraža na dva sorodna načina. Z naraščanjem dimenzije vhodnega prostora se namreč povečujeta:
  - potreba po podatkih, ki eksponentno narašča z dimenzijo in
  - število gradnikov (neuronov) ANN,

kar ima za posledico večjo računsko zahtevnost in povečanje možnosti, da se optimizacija konča v lokalnem minimumu.

Ena možnost za zmanjšanje problema prekletstva dimenzije je regularizacija parametrov [15], ki v postopku optimizacije parametrov modela “kaznuje” prevelike vrednosti parametrov, s čimer se zmanjša prenaučenosť in poveča gladkost opisovane funkcije.

## Mreža lokalnih modelov

Druga možnost za zmanjšanje prekletstva dimenzije, ki hkrati poveča transparentnost modela, je uporaba mreže lokalnih modelov (LMN) [32, 74]. S tem vedno bolj popularnim pristopom sistem namesto z enovitim *globalnim* modelom opišemo z množico oz. mrežo *lokalnih* modelov, ki opisujejo neznan sistem vsak na svojem (ožjem) področju delovanja (slika 3.1). Globalni opis sistema dobimo z zlivanjem lokalnih modelov glede na trenutno stanje sistema, ki je opisano z razvrstilnimi spremenljivkami (angl. *scheduling variables*). Te so argument utežnostnim funkcijam (angl. *validity functions*, tudi *blending functions*, *membership functions*), ki določajo pomembnost posameznega modela glede na stanje sistema. Navadno so razvrstilne spremenljivke podmnožica regresorjev modela<sup>1</sup>.



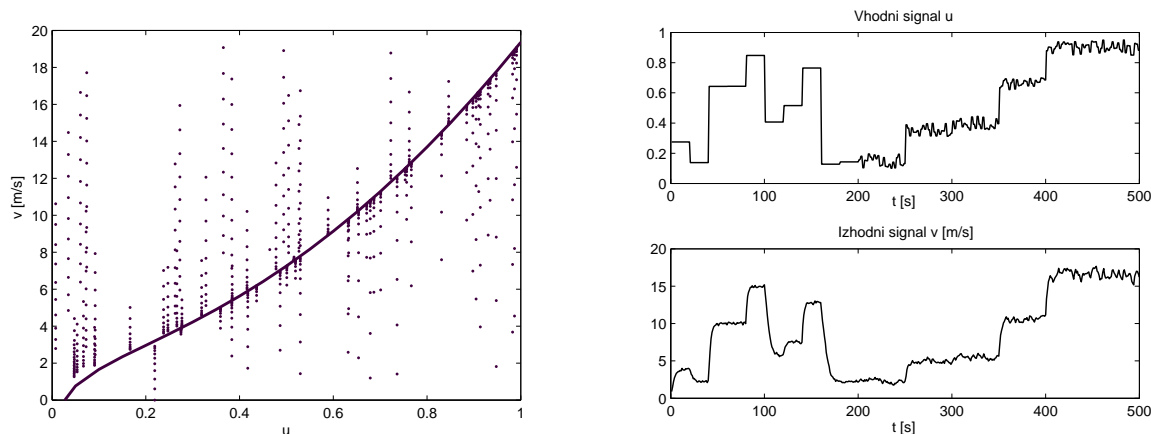
**Slika 3.1:** Modeliranje nelinearnega dinamičnega sistema z mrežo lokalnih modelov [32]. Razbijanje delovnega področja na podpodročja (levo) in prispevek posameznega lokalnega modela (desno). Sliki sta vzeti iz [32] z dovoljenjem avtorja.

Čeprav se predlagani pristop obnese bolje, kot se npr. ANN, še vedno prihaja do težav pri opisu *dinamičnih* sistemov [35, 37, 75]:

1. Problem opisovanja dinamike sistema zunaj ravnotežnih področij [37, 75] izhaja iz “težnje” stabilnih dinamičnih sistemov k ravnotežnim področjem. Sistem je zato le malo časa na območjih stran od ravnotežne krivulje in za opis teh področij imamo na voljo le malo podatkov. Ta problem je za preprost model dinamike vozila [75] orisan na sliki 3.2. Vhod v sistem  $u$

<sup>1</sup>Regresorji so stanja sistema in vhodi v sistem, ki določajo njegov odziv. Primer regresorjev v vhodno/izhodnem modelu so zakasnjene vrednosti vhodov in izhodov.

predstavlja relativni premik pedala za plin, izhod  $v$  pa je hitrost vozila. Večina vzorcev vrednosti izhodnega signala leži ob ravnotežni krivulji.



**Slika 3.2:** Problem opisovanja dinamike sistema v neravnotežnih področjih na primeru preprostega modela dinamike vozila [75], kjer je  $u$  relativni položaj pedala za plin in  $v$  hitrost vozila. Na levi sliki so s pikami označeni vzorci meritev, s polno črto pa ravnotežna krivulja. Signala, iz katerih smo podatke vzorčili, sta prikazana na desni sliki.

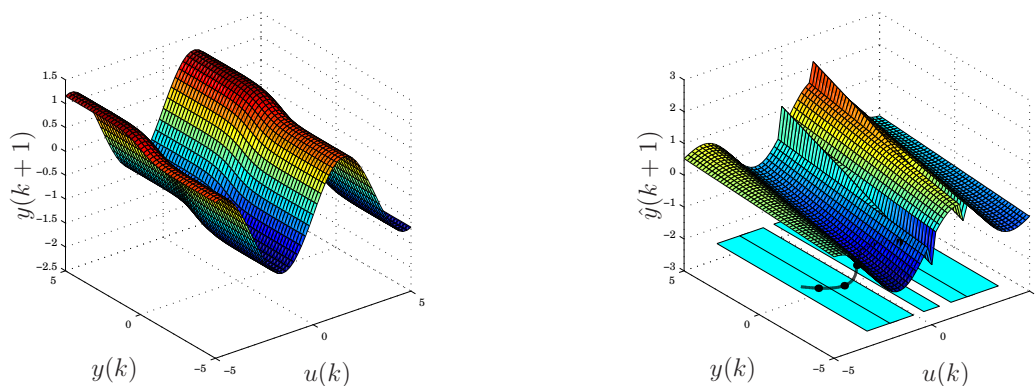
Ta problem ne bi bil tako velik, če bi v praksi sisteme lahko poljubno vzbujali in tako dobili opis sistema na celotnem delovnem območju, vendar smo zaradi narave procesa in drugih omejitev pri načrtovanju vzbujevalnega signala navadno omejeni.

2. Drugi problem je optimizacija parametrov, ta določa izbrano dinamiko, ki naj jo LMN opisuje. Na voljo imamo dva pristopa [32, 75]:
  - Prva možnost je optimizacija globalne dinamike neznanega sistema. V tem primeru optimiramo vse parametre lokalnih modelov, ki sestavljajo LMN, naenkrat, pri čemer uporabimo vse učne podatke, ki jih imamo na voljo. Optimirani parametri med seboj niso neodvisni in lokalni modeli tudi v točkah, v okolici katerih so bili identificirani, ne predstavljajo natančne linearizacije identificiranega sistema [32].
  - Druga možnost je optimizacija lokalne dinamike, pri kateri parametre optimiramo za vsak lokalni model posebej. Ena možnost je uporaba algoritma uteženih najmanjših kvadratov, druga pa ločena identifikacija lokalnih modelov [64, 69] na podatkih, ki opisujejo samo ustrezno področje sistema [32].

Postopek optimizacije globalne dinamike daje navadno globalno boljši opis sistema, ker se prosti parametri lokalnih modelov neravnotežnih področij

lahko prilagodijo tako, da povečajo veljavnost lokalnih modelov. Posledica boljšega globalnega obnašanja je izguba informacije o lokalni dinamiki, saj je parametri lokalnih modelov ne odražajo več. Če nasprotno lokalne modele optimiramo z drugim pristopom, ti predstavljajo lokalno dinamiko in so s tem bolj transparentni ter bolj uporabni za analizo in načrtovanje vodenja, a so hkrati veljavni za manjše področje delovanja. Posledično lahko nekatera področja delovanja sistema ostanejo neopisana, kar pripelje nazaj na problem opisovanja področij zunaj ravnotežne krivulje (prim. prvo točko).

3. Problemu neopisanosti nekaterih področij se lahko izognemo tudi z uporabo manjšega vektorja razvrstilnih spremenljivk [37]. Ta pristop poveča doseg veljavnosti lokalnih modelov, vendar se pojavi problem zlivanja lokalnih modelov na področjih, oddaljenih od ravnotežne krivulje. Zlivanje lahko postane negladko ali celo nezvezno [35], kar lahko privede do nestabilnosti pri vodenju [35]. Na sliki 3.3 levo je prikazana površina, ki določa nelinearnost identificiranega sistema, desno je njen opis z LMN, pri čemer je bil uporabljen zmanjšan vektor razvrstilnih spremenljivk. Vidne so nezveznosti na mejah med področji, opisanimi s posameznimi lokalnimi modeli, na področjih, ki so bolj oddaljena od ravnotežne krivulje.



**Slika 3.3:** Problem zlivanja LMN pri uporabi manjšega vektorja razvrščevalnih spremenljivk [32], originalni sistem (levo) in LMN z zmanjšanim vektorjem razvrstilnih spremenljivk (desno). S črno črto je označena ravnotežna krivulja, s točkami pa središča lokalnih modelov. Sliki sta vzeti iz [32] z dovoljenjem avtorja.

4. Še en problem je, kako ustrezno razdeliti področje delovanja med lokalne modele (angl. *partitioning problem*) oz. kako izbrati lego lokalnih modelov in velikosti področij, ki naj jih lokalni modeli opisujejo. Za rešitev tega problema se ponujata dve nasprotni metodi [32]. Metoda drobljenja področja (angl. *foreword regression*) predpostavlja preprosto začetno strukturo, ki nato raste glede na kompleksnost sistema. V vsakem koraku

se strukturi dodajo lokalni modeli, ki opisujejo preslabo opisana področja, dokler z opisom sistema nismo zadovoljni. Metoda združevanja podpodročij (angl. *backward regression*) začne z zelo kompleksno strukturo, ki pa se zmanjšuje z združevanjem tistih sosednjih lokalnih modelov, ki so si med seboj dovolj podobni. Bolje je problem izbire strukture LMN opisan v [32] in v tamkajšnjih referencah.

5. Zadnji problem, ki ga bomo samo omenili, je problem izbire reda lokalnih modelov, iz katerih želimo sestaviti LMN, in je opisan s Takensovimi teoremom (Dodatek D). Čeprav bi vsakega od identificiranih lokalnih modelov, ki opisujejo nelinearni sistem, lahko zapisali z minimalno izvedbo (načelno linearne sisteme  $d$ -tega reda lahko opišemo z linearnim modelom  $d$ -tega reda), je zaradi zlivanja posameznih identificiranih lokalnih modelov, s katerimi dobimo globalni model, potrebno vse modele opisati v isti izvedbi, ki pa je lahko neminimalna [56]. Dodatni pogoji, ki jih moramo pri tem upoštevati, so navedeni v [55].

Nekaterim naštetim problemom se lahko izognemo, če za izvedbo LMN uporabimo hitrostno linearizacijo (angl. *velocity-based linearisation*, VBL) [53, 55], ki je orodje, razvito za modele zveznih nelinearnih dinamičnih sistemov. Ideja tega pristopa je, da namesto stanja oz. izhoda sistema opisujemo njune "odvode", s čimer izničimo vpliv statične komponente stanja sistema. Na ta način delamo z linearnimi lokalnimi modeli in ne z afinimi lokalnimi modeli, ki povzročajo probleme pri opisovanju dinamike z LMN. Izvedba s hitrostno linearizacijo se dobro obnese pri hitrih spremembah delovnega območja sistema, saj ta način bolje opisuje dinamični sistem zunaj ravnotežnega področja, zaradi česar se bo njena uporaba v praktičnih aplikacijah [51, 54, 57] verjetno povečala.

Glede na sistem, ki ga želimo identificirati, so naštetih problemi bolj ali manj izraziti. Z uporabo različnih metod ali njihovih kombinacij se da te probleme zmanjšati. Da bi rešili prav te konkretne probleme, ki smo jih naštetili, je bila za identifikacijo nelinearnih dinamičnih sistemov predlagana uporaba GP modelov [35, 75].

## 3.2 Identifikacija dinamičnih sistemov z GP modelom

### 3.2.1 Lastnosti GP modela

GP model se od večine naštetih modelov razlikuje po tem, da je *neparametričen*, saj informacije o neznanem sistemu ne vsebuje samo v (hiper)parametrih kovariančne funkcije, ampak je obnašanje sistema opisano tudi z učnimi podatki  $\mathcal{D}$ , ki jih model vsebuje.

Strnjene dobre lastnosti GP modela, zaradi katerih bi ga bilo smiselno uporabiti kot orodje za identifikacijo dinamičnih sistemov, so zbrane v naslednjih alinejah:

- Napoved je dana v obliki Gaussove porazdelitve, pri čemer na varianco lahko gledamo kot na mero zaupanja v napoved ali kot na mero negotovosti. Ta je odvisna od kvalitete in lege učnih podatkov v prostoru<sup>2</sup>. Kot bo pokazano, jo lahko koristno uporabimo pri vrednotenju modela, napovedovanju, zaznavanju napak itd.
- Število hiperparametrov, ki jih moramo v postopku identifikacije optimirati, je razmeroma majhno. Tako je pri uporabi Gaussove kovariančne funkcije (2.4) in dimenziji vhodnega prostora  $D$  potrebno določiti le  $D + 2$  parametrov, medtem ko bi bilo število parametrov primerljive ANN precej večje.
- Model je precej robusten, saj v primerjavi z drugimi modeli deluje razmeroma dobro tudi z malo podatki za učenje (npr. v neravnotežnem področju [75]).
- Prekletstvo dimenzije je manj izraženo, saj: (a) GP model rabi razmeroma majhno število učnih podatkov in (b) optimirati je potrebno samo vrednosti relativno majhnega števila hiperparametrov. Kljub temu da je problem zmanjšan, je sistem z večjim številom regresorjev še vedno potrebno predstaviti s precej točkami, ki predstavljajo sistem s hiperpovršino v  $D + 1$ -dimenzionalnem prostoru.
- Model deluje razmeroma dobro tudi, ko so podatki, ki opisujejo sistem, relativno močno pošumljeni.
- Možno je vključevanje različnih vrst predznanja, ki lahko izboljša delovanje modela, pohitri optimizacijo in napovedovanje modela ali poveča njegovo transparentnost. Vključevanje predznanja bo opisano v četrtem poglavju.
- Model je enostaven za uporabo. Ne uporablja razvrstilnih spremenljivk, prav tako ni potrebna delitev delovnega področja na podpodročja v primerjavi z npr. LMN.

Seveda ima GP model tudi nekaj omejitev:

- Je neparametričen in zato neuporaben, če potrebujemo parametričnega (npr. za nekatere postopke načrtovanja vodenja).

---

<sup>2</sup>Dana "mera zaupanja" v napovedan izhod ni izključno domena GP modelov. Kako jo dobiti za mehke modele je opisano npr. v [105], za ANN pa npr. v [59].

- Navadno je manj pregleden, vendar transparentnost lahko povečamo z vključevanjem predznanja.
- Uporaba modela je računsko zahtevna v primeru velike učne množice, predvsem to velja v fazi optimizacije hiperparametrov, kar otežuje uporabo GP modela npr. za adaptivno vodenje. Računsko zahtevnost učenja lahko zmanjšamo na več načinov. Ena od možnosti je zgoščevanje učne informacije v lokalni model, kar bomo prikazali za dinamične sisteme v podpoglavju 4.2. Druga možnost so pohitritve, ki nastanejo z uporabo manjše podmnožice učne množice, prim. npr. [96], ali z aproksimacijo inverza kovariančne matrike, ki je računsko najbolj zahtevna [121]. Poleg teh in še nekaterih drugih metod, opisanih v [88, 91] in tamkajšnjih referencah, je možna tudi uporaba vzporednega računanja (angl. *parallel computing*).

Povezava med GP modeli in nekaterimi drugimi modeli je podrobneje opisana v [66, 91] in tamkajšnjih referencah, povezava med GP modelom in ANN pa v [66, 79]. Primerjava uporabe GP modela in ANN za identifikacijo dinamičnih sistemov je predstavljena v [45].

### 3.2.2 Predstavitev dinamičnih sistemov z GP modelom

Poglejmo, kako bi GP model uporabili za opis dinamičnega sistema. GP model, predstavljen v drugem poglavju, je diskretni vhodno/izhodni model. Prvotno je bil uporabljen za modeliranje statičnih sistemov:  $\mathbf{x} \rightarrow f(\mathbf{x})$ , v katerih kot vhod v model  $\mathbf{x}$  nastopajo vrednosti neodvisnih spremenljivk,  $f(\mathbf{x})$  pa je poljubna (nelinearna) funkcija, ki jo želimo opisati. Zanima nas, kako se spremeni GP model, če ga namesto za opis statičnega sistema uporabimo za opis dinamičnega sistema.

Opis vhodno/izhodne preslikave  $\mathbf{x} \rightarrow f(\mathbf{x})$  ni odvisen od dinamične/statične lastnosti sistema, toda spremeni se pomen posameznih komponent vhodnega vektorja. Pri opisovanju dinamičnega sistema vhod v GP model v nekem trenutku ni odvisen od časovno neodvisnih vrednosti vhodnih spremenljivk, marveč od preteklih vrednosti vhodov in izhodov iz sistema oz. (splošneje) stanja sistema. Najbolj tipično nelinearni dinamični sistem opišemo z nelinearnim regresijskim modelom s posplošenim pogreškom, znanim pod kratico NARX (angl. *Nonlinear AutoRegressive model with eXogenous input*) [97]; pri tem model dinamičnega sistema  $L$ -tega reda za napovedovanje izhoda  $\hat{y}(k)$  uporablja vhodni vektor  $\mathbf{x}_k$ , v katerem nastopajo pretekle vrednosti vhoda v sistem  $u$  in izhoda iz sistema  $y$ :

$$\hat{y}(k) = f(\mathbf{x}_k) \quad (3.1)$$

$$\mathbf{x}_k = [y(k-1) \dots y(k-L) \ u(k-1) \dots u(k-L)]. \quad (3.2)$$

Tako je dinamični sistem navadno opisan tudi, ko ga identificiramo z ANN.

\*       \*       \*

Kdaj torej uporabiti GP model za identifikacijo dinamičnih sistemov namesto bolj standardnih metod?

- Informacija o sistemu je dana v obliki vhodno/izhodnih podatkov.
- Podatki, ki so na voljo, so lahko slabši – veliko šuma, napake meritev.
- Potrebujemo ali želimo mero negotovosti v napovedi.
- Na voljo imamo razmeroma majhno število podatkov glede na število regresorjev.

Oglejmo si postopek identifikacije dinamičnega sistema z GP modelom.

### 3.3 Postopek identifikacije

Predstavili bomo postopek identifikacije dinamičnih sistemov z GP modelom, ki smo ga na kratko predstavili že v [10, 12]. Pri identifikaciji modela sistema ne dobimo iz teoretičnega opisa obravnavanega sistema, marveč iz meritev, seveda pa si pri tem lahko pomagamo z znanjem o sistemu, če ga imamo. O vključevanju predznanja v GP model bomo sicer govorili v četrtem poglavju, tu bo predznanje uporabljeno samo posredno kot pomoč za identifikacijo. V grobem je identifikacija z GP modelom sestavljena iz naslednjih faz:

1. definicija namena modela;
2. postavitve GP modela;
3. načrtovanje eksperimenta z uporabo a priori znanja ali predhodnih meritev;
4. eksperiment in obdelava signalov, da pridobimo podatke za učenje in vrednotenje;
5. učenje GP modela, tj. optimizacija hiperparametrov, in
6. vrednotenje GP modela.

V vsaki fazi se lahko odločimo, ali bomo nadaljevali z naslednjo fazo identifikacije ali se bomo vrnili na katero izmed prejšnjih. Celoten postopek identifikacije je redko opravljen v enem koraku, ampak je ponavadi iterativen. Identifikacija je končana, ko se na podlagi vrednotenja odločimo, da je model dovolj dober za svoj namen. Shema postopka identifikacije je prikazana na sliki 3.4.

V nadaljevanju bomo najprej na kratko predstavili posamezne faze identifikacije. Večji del faz identifikacijskega postopka je enak kot pri drugih identifikacijskih postopkih [64, 69]. Poudarili bomo tiste faze identifikacijskega postopka, ki se zaradi lastnosti GP modela bolj razlikujejo od istih faz pri drugih vrstah modelov.

### 3.3.1 Definicija namena modela

Namen identifikacije je dobiti model nekega procesa. Ta model ni sam sebi namen, ampak naj bi bil namenjen neki nalogi. Ta naloga določa kriterije, ki jih mora model izpolniti pri vrednotenju, da ga sprejmemo kot zadovoljiv model. Model uporabljamo za simulacijo, napovedovanje odziva, načrtovanje vodenja, analizo, zaznavanje napak ipd.

### 3.3.2 Postavitev modela

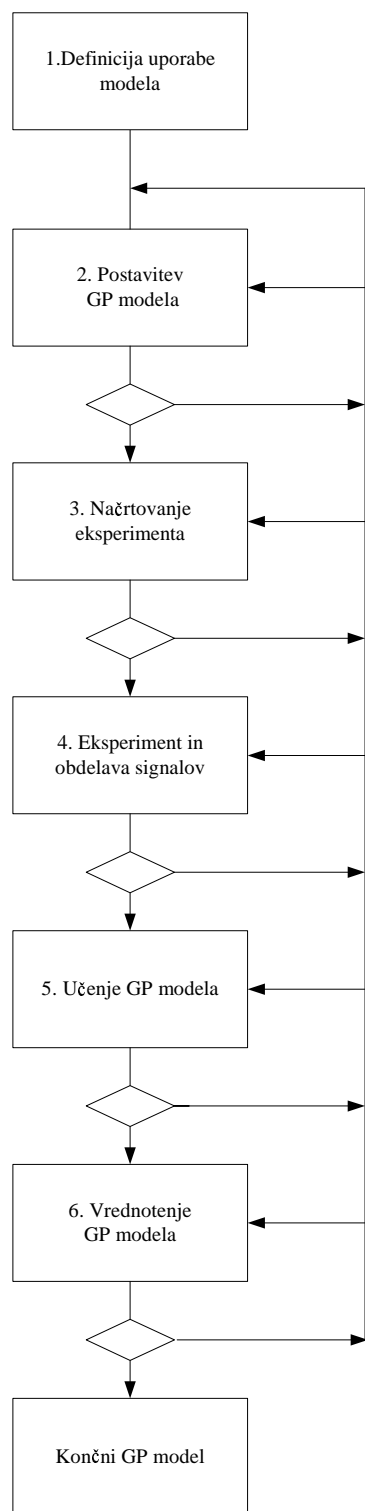
Naslednja faza v identifikaciji je izbira modela in njegovih elementov, pri čemer se bomo osredotočili na uporabo GP modela. Na podlagi a priori znanja o procesu, morebitnih predhodnih meritev ali *rezultatov vrednotenja* postavimo GP model. Postavitev obsega:

- izbiro primerne kovariančne funkcije,
- izbiro regresorjev in
- odločitev o morebitnem vključevanju predznanja.

Pri dinamičnih sistemih z izbiro regresorjev določimo tudi red modela. O tem problemu bo več govora v podpoglavju 3.4 o postavitvi GP modela, kjer bomo podrobneje predstavili to zelo pomembno fazo identifikacijskega postopka.

### 3.3.3 Načrtovanje eksperimenta

Naslednja faza v postopku identifikacije je načrtovanje eksperimenta. Tega načrtujemo na podlagi znanja o procesu, dobljenega iz predznanja ali predhodnih meritev, in glede na postavljeni GP model.



Slika 3.4: Potek identifikacije

Postopek načrtovanja je soroden kot pri drugih metodah:

1. izbira vhodov in izhodov iz procesa, pri čemer moramo paziti, da merimo vse vplivne veličine;
2. izbira primerne časa vzorčenja;
3. izbira primerne vzbujevalnega signala, da je proces opisan na celotnem področju delovanja, ki nas zanima. Lastnost GP modela je namreč, da model dobro interpolira na področju, na katerem imamo učne podatke, in slabo ekstrapolira.

Pri načrtovanju moramo upoštevati omejitve, kot so:

- motnje merjenih signalov;
- omejitev vhodnega signala v proces zaradi fizičnih ali varnostnih razlogov in omejitve izvršnih členov;
- omejen čas, ki je na voljo za izvedbo identifikacijskega postopka (recimo pri merjenjih v industrijskem okolju);
- težave pri zajemanju podatkov v zaprti zanki.

### 3.3.4 Eksperiment in obdelava podatkov

Eksperiment izvedemo, kot smo ga načrtali v prejšnjem razdelku. Rezultat eksperimenta so izmerjeni vhodni in izhodni signali procesa, ki predstavljajo njegovo obnašanje.

Vhod v GP niso signali, marveč vzorci, ki določajo obnašanje sistema/procesa pri določenih vrednostih regresorjev v izbranem trenutku. Te vzorce dobimo z vzorčenjem vhodnih in izhodnih signalov procesa, pri čemer čas vzorčenja določimo že v fazi načrtovanja eksperimenta.

Primer možne izbire vhodno/izhodnega učnega para (z indeksom  $i$ ) za proces z vhodnim signalom  $u$  in izhodnim signalom  $y$ , ki bi ga radi opisali s sistemom drugega reda:

- učni izhod  $i$ : izhod iz procesa v času  $t = kT$ :  $y_i = y(k)$ ;
- ustrežni učni vhod:  $\mathbf{x}_i = [y(k-1) \ y(k-2) \ u(k-1) \ u(k-2)]$ ,

kjer je  $T$  čas vzorčenja vhodnih in izhodnih signalov. Testne vhode sestavljamo enako kot učne.

Dobra lastnost GP modela je, da se pri določeni izbiri kovariančne funkcije pomembnost posameznih regresorjev odraža skozi vrednost ustreznih hiperparametrov (angl. *automatic relevance determination*, ARD) [79]. Za Gaussovo kovariančno funkcijo (2.4) npr. velja, da večja ko je vrednost hiperparametra  $w_d$ , pomembnejši je ustrezni regresor  $d$ . To lastnost lahko uporabimo samo, če dobljene učne vzorce pred uporabo za učenje *normaliziramo* tako, da so vzorci, ki predstavljajo posamezni regresor, usrediščeni s približno enakim odstopanjem od srednje vrednosti, npr. vse vrednosti vzorcev nekega regresorja so med -1 in 1. Drugi pogoj je, da smo optimizacijo končali v globalnem minimumu.

V primeru manjkajočih vrednosti regresorjev v posameznem učnem vzorcu lahko te vrednosti, ali nadomestimo z nekim (primernim) povprečjem prejšnjih in poznejših vzorcev (GP model povpreči informacijo) ali pa tak vzorec zavržemo ali zamenjamo z drugim, če imamo na voljo dovolj podatkov.

### 3.3.5 Učenje modela

Učenje GP modela je enako, naj gre za GP model, ki opisuje statične sisteme ali GP model, ki opisuje dinamične. V postopku učenja optimiramo hiperparametre  $\Theta$ . Ti niso znani vnaprej, ampak jih je potrebno določiti iz učnih podatkov. Običajno je uporabljena metoda največje podobnosti, saj kljub preprostosti daje dobre rezultate. Za določitev najbolj verjetnih vrednosti hiperparametrov smo uporabljali metodo, ki temelji na Polack-Ribiereovi metodi konjugiranih gradientov C. E. Rasmussena<sup>3</sup>, v splošnem pa bi za to nalogo lahko uporabljali katerokoli optimizacijsko metodo [27].

Zaradi možnosti, da postopek optimizacije obtiči v lokalnem minimumu, učenje navadno ponovimo večkrat za različne začetne vrednosti hiperparametrov in vrednotimo tako dobljene modele.

### 3.3.6 Vrednotenje modela

Čeprav je vrednotenje zelo pomemben korak v identifikacijskem postopku, ki pove kako dober je dobljeni model, mu je dostikrat namenjeno premalo pozornosti. Z vrednotenjem preverimo ujemanje matematičnega modela in obravnavanega sistema [72].

<sup>3</sup>Funkcija `minimize`, del programskega paketa GPML, <http://www.gaussianprocess.org/gpml/>

Kvaliteto modela lahko merimo na več načinov, najbolj pomembni vidiki pa so:

- verodostojnost modela (angl. *model plausibility*), pri čemer nas zanima skladnost modela s predhodnim znanjem;
- (ne)veljavnost modela (angl. *model falseness*), pri tem nas navadno zanima ujemanje izhodov modela in procesa, pri čemer kot vhod uporabimo množico podatkov, ki ni bila uporabljena za modeliranje, ter
- namembnost modela (angl. *model purposiveness*), pri čemer preverimo, ali je model ustrezen za namenjeno nalogo.

Pregled metod vrednotenja najdemo v [36,72], podrobneje bodo (posebno za GP model) predstavljene v podpoglavju 3.6.

Zelo pogosto pri vrednotenju uporabljamo rezultate simulacije ali enokoračne predikcije modela, ki ju primerjamo z obnašanjem procesa, zato bo pred podrobnejšo predstavitev vrednotenja v podpoglavju 3.6 predstavljena simulacija GP modela v podpoglavju 3.5.

\* \* \*

Na kratko smo opisali vse faze identifikacije GP modela. V nadaljevanju sledi podrobnejša predstavitev tistih delov identifikacijskega postopka, pri katerih se identifikacija z GP modelom bolj loči od identifikacijskih metod z drugimi modeli ali pri katerih je kakšna od lastnosti GP modela še posebej uporabna:

- postavitvev GP modela,
- simulacija GP modela, ki bo potrebna pri vrednotenju in uporabi GP modela, in
- vrednotenje GP modela.

Potek identifikacijskega postopka bomo prikazali na primeru identifikacije modela bioreaktorja v podpoglavju 3.7 in na primeru identifikacije modela nitrifikacijskega procesa čistilne naprave v podpoglavju 5.3.

### 3.4 Postavitev GP modela

Postavitev GP modela je prvi korak v njegovi gradnji. Dva nujna elementa postavitve sta:

- izbira primerne kovariančne funkcije in
- izbira regresorjev GP modela, kar posledično določa tudi red modela.

Pri obeh elementih si pomagamo s predznanjem o sistemu in z rezultati vrednotenja predhodnih modelov. Določene oblike predznanja lahko tudi eksplicitno vključimo v GP model. Ker je postopek identifikacije za GP modele z vključenim predznanjem in brez njega podoben in je področje vključevanja predznanja v GP model obsežno, ga bomo predstavili v četrtem poglavju, skupaj s spremembami postopka identifikacije, ki nastanejo zaradi vključevanja.

#### 3.4.1 Izbira kovariančne funkcije

Izbira kovariančne funkcije  $C(\mathbf{x}_i, \mathbf{x}_j)$  je ključnega pomena za delovanje GP modela. Z njo izrazimo znanje o obnašanju oz. splošnih lastnostih neznanega sistema, naj je to a priori ali pridobljeno s predhodnim eksperimentiranjem. Edini pogoj, da je neka funkcija  $C(.,.)$  lahko kovariančna funkcija, je, da tvori pozitivno definitno kovariančno matriko  $\mathbf{K}$  (2.1) za katerokoli število  $N$  poljubnih vhodnih vektorjev  $\mathbf{x}_i$ ,  $i = 1, \dots, N$  [26].

Kovariančna funkcija  $C(\mathbf{x}_i, \mathbf{x}_j)$  izraža mero podobnosti med vhodoma  $\mathbf{x}_i$  in  $\mathbf{x}_j$ . Za realne procese je navadno sestavljena iz dveh delov:

$$C(\mathbf{x}_i, \mathbf{x}_j) = C_f(\mathbf{x}_i, \mathbf{x}_j) + C_n(\mathbf{x}_i, \mathbf{x}_j). \quad (3.3)$$

Prvi, tj. *funkcijski* del  $C_f(.,.)$ , opisuje lastnosti neznanega sistema, drugi, *šumni* del  $C_n(.,.)$ , pa predstavlja prispevek šuma. Ker je vsota dveh pozitivno definitnih matrik tudi pozitivno definitna matrika, lahko izberemo funkciji  $C_n(.,.)$  in  $C_f(.,.)$  ločeno.

V literaturi, ki opisuje GP model, npr. [26, 91], se največkrat uporablja delitev (funkcijskih delov) kovariančnih funkcij na dva razreda:

- stacionarne kovariančne funkcije in
- nestacionarne kovariančne funkcije.

*Stacionarne* kovariančne funkcije so tiste, pri katerih je vrednost funkcije  $C_f(\mathbf{x}_i, \mathbf{x}_j)$  odvisna samo od relativne lege vhodnih vektorjev  $\mathbf{x}_i$  in  $\mathbf{x}_j$  oziroma od njune medsebojne razdalje:  $C_f(\mathbf{x}_i, \mathbf{x}_i + \Delta\mathbf{x}) = C_f(\Delta\mathbf{x})$ . Primera stacionarnih funkcij sta:

- Gaussova kovariančna funkcija (2.4):

$$C_f(\mathbf{x}_i, \mathbf{x}_j) = C_G(\mathbf{x}_i, \mathbf{x}_j) = v \exp \left[ -\frac{1}{2} \sum_{d=1}^D w_d (x_i^d - x_j^d)^2 \right], \quad (3.4)$$

ki jo izberemo, če predpostavimo gladkost izhoda in stacionarnost (poglavje 2.1, str. 11), in

- funkcija

$$C_f(\mathbf{x}_i, \mathbf{x}_j) = v \exp \left[ -\frac{1}{2} \sum_{d=1}^D \left( \frac{1}{w_d} \sin^2 \left( \frac{\pi}{\lambda_d} (x_i^d - x_j^d) \right) \right) \right], \quad (3.5)$$

ki jo uporabimo, če bi radi opisali periodični proces z znanimi valovnimi dolžinami  $\lambda_d$  v smeri  $d$ -te komponente vhodnega vektorja  $\mathbf{x}$  [26]. Taka funkcija izraža veliko korelacijo ne samo med vhodi, ki ležijo blizu skupaj, marveč tudi med vhodi, ki so med seboj v smeri komponente  $d$  oddaljeni za mnogokratnike valovnih dolžin  $\lambda_d$ .

*Nestacionarne* kovariančne funkcije so tiste, pri katerih je vrednost funkcije  $C_f(\mathbf{x}_i, \mathbf{x}_j)$  odvisna od absolutne lege vhodnih vektorjev  $\mathbf{x}_i$  in  $\mathbf{x}_j$  v prostoru. Če domnevamo, da obstaja neki linearni trend v podatkih, potem lahko Gaussovi kovariančni funkciji  $C_G$  (2.4) dodamo člen [26]:

$$C_f(\mathbf{x}_i, \mathbf{x}_j) = C_G(\mathbf{x}_i, \mathbf{x}_j) + \sum_{d=1}^D \sigma_a^2 x_i^d x_j^d + \sigma_c^2, \quad (3.6)$$

ki napovedim dodaja še linearno komponento.

Največkrat, zlasti kadar ne poznamo ozadja procesa, za funkcijski del kovariančne funkcije uporabimo kar Gaussovo kovariančno funkcijo (2.4), saj sta gladkost izhoda in stacionarnost mnogokrat predpostavljene lastnosti. Primer uporabe drugačne kovariančne funkcije je za modeliranje vetrne turbine prikazan v [58].

Pri izbiri *šumnega* dela kovariančne funkcije  $C_n$  pogosto za sistem predpostavimo kar beli šum na izhodu. S tem za šumni del kovariančne funkcije dobimo:

$$C_n(\mathbf{x}_i, \mathbf{x}_j) = v_0 \delta_{ij}. \quad (3.7)$$

V primeru, da je izhod sistema moten z barvnim šumom, ustrezno spremenimo šumni del kovariančne matrike, kot bomo pokazali v podpoglavju 4.4.3 o vključevanju predznanja o barvnem šumu v GP model. Obstajajo tudi druge možnosti modeliranja šuma, povzete v [91]. Tako je npr. vključevanje modela Gaussovega šuma, pri katerem je šum odvisen od vhodnega vektorja  $\mathbf{x}$ , predstavljeno v [29], vključevanje šuma s Studentovo  $t$ -distribucijo pa v [80].

Podrobnejši opis kovariančnih funkcij in napotke za njihovo izbiro lahko najdemo v [26, 91].

### 3.4.2 Regresorji

Drugi del postavitve GP modela predstavlja izbira primernih regresorjev. Pri izbiri regresorjev si pomagamo z znanjem o sistemu, naj bo to apriorno ali dobljeno z eksperimentiranjem. Problem izbire regresorjev je skupen vsem metodam za identifikacijo nelinearnih dinamičnih sistemov. Pregled pristopov je dan v [63].

Najbolj pogost pristop k izbiri je ti. izbira regresorjev na podlagi vrednotenja (angl. *validation based regressor selection*) [63]. Iskanje “najboljše” množice regresorjev se začne iz neke osnovne množice regresorjev. Po optimizaciji in navzkrižnem vrednotenju dobljenega modela se zavržejo nepomembni ali malo pomembni regresorji in dodajo potencialno koristni. Če iščemo najboljše regresorje z več različnimi GP modeli naenkrat, se potencialno dobri modeli obdržijo, slabi pa zavržejo.

Pri določanju regresorjev GP modela si lahko pomagamo tudi z ARD lastnostjo modela, saj se v primeru normaliziranih vhodov v GP model (razdelek 3.3.4) vpliv posameznih regresorjev izraža skozi vrednosti ustreznih hiperparametrov.

Z izbiro regresorjev GP modela, ki opisuje dinamični sistem, hkrati določimo tudi red modela. Pri tem si lahko delno pomagamo z interpretacijo Takensovega teorema [86], prim. Dodatek D, ki pravi:

$$m > 2d + 1, \tag{3.8}$$

kjer je  $d$  red nelinearnega dinamičnega sistema v zveznem prostoru in je  $m$  zadostni red diskretnega modela, da lahko model sistem opiše zadovoljivo iz vhodno/izhodih podatkov. V posameznih primerih se zgodi, da dobimo zadovoljive rezultate, tudi če izberemo red modela manjši od  $m$  [32].

### 3.5 Simulacija GP modelov dinamičnih sistemov

V tem podpoglavju bomo pokazali, kako se simulira GP model dinamičnega sistema. Simulacijo lahko uporabljamo pri vrednotenju modela ali za napovedovanje obnašanja identificiranega sistema.

Nasplošno imamo za simulacijo ali večkoračno predikcijo diskretnih modelov dve možnosti:

- *direktno metodo*, pri kateri naučimo več modelov za vsak želeni horizont  $h$  posebej ali
- *iterativno metodo*, pri kateri naredimo model za enokoračno predikcijo, ki jo ponavljamo  $h$ -krat.

Problem direktne metode je, da moramo vnaprej izbrati horizont, in če ga spremenimo, moramo model učiti znova. Drug problem te metode je, da za močno nelinearne sisteme in velik horizont potrebuje na voljo veliko število podatkov [27].

Nasprotno je model za enokoračno predikcijo navadno lahko narediti, z iteracijo enokoračne predikcije pa je mogoče dobiti model za katerikoli horizont hočemo, prim. [27] in tamkajšnje reference. To metodo uporabljamo npr. pri simulaciji sistemov z ANN. Problem iterativne metode simulacije je akumulacija napake, ko se pomikamo naprej po času. Ena možnost za rešitev problema je eliminacija sistematičnih napak, ki nastanejo zaradi zaporednih enokoračnih napovedi [38, 98]. Z uporabo GP modela nasprotno dobimo tudi možnost, da napake ne odpravljamo, ampak jo z uporabo variance napovedi raje ovrednotimo [27, 47], kar bo predstavljeno v podpoglavjih o propagaciji negotovosti.

V razdelku 3.2.2 smo pokazali, kako GP model  $L$ -tega reda uporabimo za opis dinamičnega sistema z vhomom  $u$  in izhodom  $y$ , enačbi (3.1) in (3.2):

$$\begin{aligned}\hat{y}(k) &= f(\mathbf{x}_k) \\ \mathbf{x}_k &= [y(k-1) \dots y(k-L) \ u(k-1) \dots u(k-L)].\end{aligned}$$

S tem modelom lahko napovemo izhod sistema  $y(k)$  v koraku  $k$  iz vrednosti regresorjev modela  $\mathbf{x}_k$ . Tu nas bo zanimalo, kako bi tak model simulirali.

#### Postopek simulacije

GP model simuliramo podobno kot ANN. Predpostavimo, da poznamo zgodovino obnašanja dinamičnega sistema reda  $L$  do koraka  $k$ . Potem v koraku  $k+1$

poznamo celoten vhodni vektor GP modela:

$$\mathbf{x}_{k+1} = [y(k) \dots y(k-L+1) \quad u(k) \dots u(k-L+1)].$$

V koraku  $k+2$  v vhodnem vektorju  $\mathbf{x}_{k+2}$

$$\mathbf{x}_{k+2} = [y(k+1) \dots y(k-L+1) \quad u(k+1) \dots u(k-L+1)] \quad (3.9)$$

nastopa vrednost izhoda iz procesa  $y(k+1)$  v času  $k+1$ . Ker te vrednosti ne poznamo, podobno kot pri ANN, namesto prave uporabimo z GP modelom napovedano vrednost GP modela  $\hat{y}(k+1) = f(\mathbf{x}_k)$ .

Za razliko od ANN napovedana vrednost izhoda GP modela ni neka natančna vrednost, marveč porazdelitev. Tako imamo dve možnosti simulacije:

- *simulacija brez propagacije negotovosti* ali “naivna” metoda, pri kateri podobno kot pri ANN za vrednost izhoda uporabimo samo srednjo vrednost napovedi, in
- *simulacija s propagacijo negotovosti*, pri kateri uporabimo celotno napovedano porazdelitev.

### 3.5.1 Simulacija brez propagacije negotovosti

V tem razdelku si bomo ogledali simulacijo brez propagacije negotovosti. Enako kot prej predpostavimo, da poznamo vse vrednosti vhoda  $u$  in izhoda  $y$  do koraka  $k$ , medtem ko naprej poznamo samo vrednosti vhoda  $u$ . Postopamo tako, da nepoznane vrednosti izhoda  $y(k+n)$  nadomestimo s srednjo vrednostjo ocene GP modela  $m(\mathbf{x}_{k+n})$ .

*Postopek simulacije*

- Korak  $k+1$ :  
Vhodni vektor:

$$\mathbf{x}_{k+1} = [y(k) \dots y(k-L+1) \quad u(k) \dots u(k-L+1)].$$

Izhod iz GP modela je:

$$\hat{y}(k+1) = \mathcal{N}(m(\mathbf{x}_{k+1}), v(\mathbf{x}_{k+1})),$$

pri tem bomo z  $v(\mathbf{x}_{k+n})$  označevali varianco napovedi izhoda sistema v  $k+n$ -tem koraku.

- Korak  $k + 2$ :  
Kot novi vhodni vektor v GP model nastopa:

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

v katerem ne poznamo vrednosti izhoda v koraku  $k + 1$ . Nadomestimo ga s srednjo vrednostjo napovedi  $m(\mathbf{x}_{k+1})$ :

$$\mathbf{x}_{k+2} = [m(\mathbf{x}_{k+1}) \ y(k) \ \dots \ y(k-L+2) \ u(k+1) \ \dots \ u(k-L+2)].$$

Izhod iz GP modela je:

$$\hat{y}(k+2) = \mathcal{N}(m(\mathbf{x}_{k+2}), v(\mathbf{x}_{k+2})).$$

- $\vdots$

- Korak  $k + n$ :  
Vhodni vektor:

$$\mathbf{x}_{k+n} = [m(\mathbf{x}_{k+n-1}) \ \dots \ m(\mathbf{x}_{k+n-L}) \ u(k+n-1) \ \dots \ u(k+n-L)]. \quad (3.10)$$

Izhod GP modela:

$$\hat{y}(k+n) = \mathcal{N}(m(\mathbf{x}_{k+n}), v(\mathbf{x}_{k+n})). \quad (3.11)$$

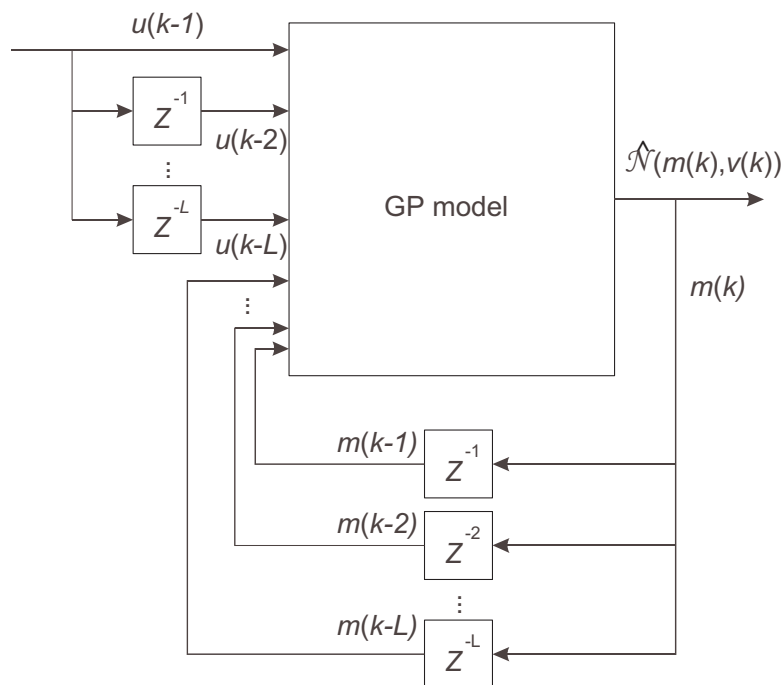
- $\vdots$

Postopek ponavljamo do zelenega končnega časa simulacije. Ilustracija tega postopka je na sliki 3.5.

Prednosti “naivnega” pristopa pred pristopom s propagacijo negotovosti sta:

- enostavnost in
- hitrost.

Njegova slabost je, da je ocenjeno zaupanje v napoved izhoda v koraku  $k + n$ , ki se izraža skozi varianco  $v(\mathbf{x}_{k+n})$ , preveliko, saj izhaja iz predpostavke o *natančni* vrednosti vhodnega vektorja  $\mathbf{x}_{k+n}$ . Ker pri “naivnem” pristopu kot vhod v GP model nastopajo samo srednje vrednosti napovedanih preteklih izhodov  $m(\mathbf{x}_{k-l})$ ,  $l = 1, \dots, L$ , je napovedana izhodna porazdelitev ožja, kot bi bila v primeru, da bi kot vhod v model uporabili celotne napovedane porazdelitve  $\hat{y}(k+n-l)$ . Kljub vsemu daje postopek še vedno razmeroma uporabno oceno, saj vsebuje informacijo o zaupanju modela na področju, na katerem je.



**Slika 3.5:** Simulacija nelinearnega dinamičnega sistema z GP modelom brez propagacije negotovosti

### 3.5.2 Simulacija s propagacijo negotovosti

Druga možnost simulacije je z uporabo napovedane variance, lastnosti, ki jo večina drugih modelov nima. Postopek je podrobno opisan v [27], skrajšana različica pa je v [28, 47] in [4].

V prejšnjem razdelku smo si ogledali potek “naivne” simulacije za čase, ko ne poznamo več izhodov iz sistema in tako nimamo več podatkov o pravih vseh v GP model. Namesto pravih vrednosti so bile uporabljene samo srednje vrednosti napovedi. Če nas zanimajo natančnejše napovedi, lahko za vhod v GP model uporabimo celotne napovedane izhodne porazdelitve GP modela. V  $(k + n)$ -tem koraku simulacije tako ne uporabimo vhodnega vektorja (3.10), ampak za  $\hat{y}(k + n - l)$ ,  $l = 1, \dots, L$ , namesto srednjih vrednosti  $m(\mathbf{x}_{k+n-l})$  uporabimo napovedane porazdelitve  $\mathcal{N}(m(\mathbf{x}_{k+n-l}), v(\mathbf{x}_{k+n-l}))$ .

Kaj se zgodi z napovedovanjem z uporabo GP modela, ko na njegovem vhodu nastopajo porazdeljene (verjetnostne) vrednosti namesto natančnih? Spomnimo se opisa sistema v poglavju 2.1, le da obravnavamo sistem brez šuma na izhodu:  $v_0 = 0$ .

Pri novem vhodu  $\mathbf{x}$  (tam označenim z  $\mathbf{x}^*$ , tu bomo  $*$  zaradi enostavnosti izpuščali) dobimo kot izhod Gaussovo porazdelitev  $p(y|\mathcal{D}, \mathbf{x}) = \mathcal{N}(\mu(\mathbf{x}), \sigma^2(\mathbf{x}))$ , katere

srednja vrednost in varianca sta dani z:

$$\mu(\mathbf{x}) = \sum_{i=1}^N \beta_i C(\mathbf{x}_i) \quad (3.12)$$

$$\sigma^2(\mathbf{x}) = C(\mathbf{x}, \mathbf{x}) - \sum_{i,j=1}^N K_{ij}^{-1} C(\mathbf{x}, \mathbf{x}_i) C(\mathbf{x}, \mathbf{x}_j), \quad (3.13)$$

kjer je  $\boldsymbol{\beta} = \mathbf{K}^{-1}\mathbf{y}$ .

Predpostavimo, da testni vhodni vektor ni dan kot neka točka, ampak so vrednosti posameznih regresorjev porazdeljene po Gaussovi porazdelitvi:  $\mathbf{x} = \boldsymbol{\nu} + \boldsymbol{\epsilon}_x$ ,  $\boldsymbol{\epsilon}_x \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_x)$ . Ker bi radi dobili napoved pri porazdeljenem vходу  $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$ , moramo integrirati napovedano porazdelitev  $p(y|\mathcal{D}, \mathbf{x})$  pri natančnem vходу  $\mathbf{x}$  prek vhodne porazdelitve:

$$p(y|\mathcal{D}, \boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = \int_{-\infty}^{\infty} p(y|\mathcal{D}, \mathbf{x}) p(\mathbf{x}|\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) d\mathbf{x}, \quad (3.14)$$

kjer je  $y$  normalno porazdeljen:  $p(y|\mathcal{D}, \mathbf{x}) = \frac{1}{\sqrt{2\pi\sigma^2(\mathbf{x})}} \exp\left[-\frac{1}{2} \frac{(y-\mu(\mathbf{x}))^2}{\sigma^2(\mathbf{x})}\right]$ . Ker je  $p(y|\mathcal{D}, \mathbf{x})$  nelinearna funkcija  $\mathbf{x}$ -a, nova napovedana porazdelitev  $p(y|\mathcal{D}, \boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$  ni Gaussova in integrala ne moremo rešiti, ne da bi se zatekli k aproksimaciji [27].

Imamo dve možnosti aproksimacije [27, 66]:

1. analitično aproksimacijo integrala, pri kateri izhodno *ne-Gaussovo porazdelitev* aproksimiramo z Gaussovo z enako srednjo vrednostjo in varianco ter
2. numerične Monte-Carlo (MCMC) metode.

Propagacija negotovosti z analitično aproksimacijo pri predpostavki o Gaussovi porazdelitvi izhoda je opisana v Dodatku B, propagacija negotovosti z numerično aproksimacijo integrala (3.14), ki jo bomo uporabili v podpoglavju 5.6, pa v naslednjem razdelku.

### 3.5.3 Numerična propagacija negotovosti z uporabo MCMC metod

V prejšnjem podpoglavju smo si ogledali možnosti aproksimacije pri simulaciji GP modela s propagacijo negotovosti. Druga našeta možnost je bila numerična propagacija negotovosti z uporabo MCMC (angl. *Markov Chain Monte Carlo*) metod [66]. Podobno kot pri analitični metodi aproksimacije, prim.

Dodatek B, tudi tu aproksimiramo integral (3.14), le da tu na rezultat sklepamo s “povprečenjem” rezultatov, ki jih daje velikokrat ponovljena simulacija.

Enako kot pri analitični metodi propagacije negotovosti je tudi pri tej metodi vhod v GP model *napovedana izhodna porazdelitev*. Recimo, da poznamo vrednosti izhoda iz sistema do časa  $k$ . Kot vemo, je v koraku  $k + 1$  izhodna porazdelitev Gaussova:  $\hat{y}(k + 1) = \mathcal{N}(\mu(\mathbf{x}_{k+1}), \sigma^2(\mathbf{x}_{k+1}))$ . V koraku  $k + 2$  je vrednost izhoda GP modela po enačbi (3.14):

$$p(y_{k+2} | \mathcal{D}, \boldsymbol{\nu}_{k+2}, \boldsymbol{\Sigma}_{\mathbf{x}_{k+2}}) = \int_{-\infty}^{\infty} p(y_{k+2} | \mathcal{D}, \mathbf{x}_{k+2}) p(\mathbf{x}_{k+2}) d\mathbf{x}_{k+2}. \quad (3.15)$$

Namesto analitične aproksimacije ta integral aproksimiramo numerično z MCMC metodo:

$$p(y_{k+2} | \mathcal{D}, \boldsymbol{\nu}_{k+2}, \boldsymbol{\Sigma}_{\mathbf{x}_{k+2}}) = \frac{1}{S} \sum_{s=1}^S p(y_{k+2} | \mathcal{D}, \mathbf{x}_{k+2}^s), \quad (3.16)$$

kjer je  $S$  število vzorcev in je  $\mathbf{x}_{k+2}^s$  vzorec iz porazdelitve  $p(\mathbf{x}_{k+2})$ . Izhodna porazdelitev je torej mešanica Gaussovih porazdelitev:

$$p(y_{k+2} | \mathcal{D}, \boldsymbol{\nu}_{k+2}, \boldsymbol{\Sigma}_{\mathbf{x}_{k+2}}) = \frac{1}{S} \sum_{s=1}^S \mathcal{N}(\mu(\mathbf{x}_{k+2}^s), \sigma^2(\mathbf{x}_{k+2}^s)). \quad (3.17)$$

Sedaj se pojavi problem, saj izhod  $\hat{y}(k + 2)$  ni v obliki Gaussove porazdelitve, ampak je mešanica Gaussovih porazdelitev (angl. *Gaussian mixtures*). Tako moramo v tem in vsakem naslednjem koraku vzorčiti vedno bolj zapleteno mešanico Gaussov, kar je za večje število korakov računsko zelo zamudno. Namesto tega se zatečemo k aproksimaciji opisanega postopka [27]:

### Numerična aproksimacija: prvi postopek

- zanka po  $s$

–  $k + 1$ :

$$\mathbf{x}_{k+1} = [y(k)y(k-1) \dots y(k-L)]$$

izračunaj  $\hat{y}(k+1) \sim \mathcal{N}(\boldsymbol{\nu}_{k+1}, \boldsymbol{\Sigma}_{k+1})$

–  $k + 2$ :

$$\mathbf{x}_{k+2} = [\hat{y}(k+1)y(k) \dots y(k-L+1)] \sim \mathcal{N}(\boldsymbol{\nu}_{k+2}, \boldsymbol{\Sigma}_{k+2})$$

vzorči  $\mathbf{x}_{k+2}^s$  iz  $p(\mathbf{x}_{k+2})$

izračunaj  $p(y(k+2) | \mathcal{D}, \boldsymbol{\nu}_{k+2}, \boldsymbol{\Sigma}_{k+2}) = \mathcal{N}(\mu(\mathbf{x}_{k+2}^s), \sigma^2(\mathbf{x}_{k+2}^s))$

–  $k + 3$ :

$$\mathbf{x}_{k+3} = [\hat{y}(k+2)\hat{y}(k+1) \dots y(k-L+2)] \sim \mathcal{N}(\boldsymbol{\nu}_{k+3}, \boldsymbol{\Sigma}_{k+3})$$

vzorči  $\mathbf{x}_{k+3}^s$  iz  $p(\mathbf{x}_{k+3})$

izračunaj  $p(y(k+3) | \mathcal{D}, \boldsymbol{\nu}_{k+3}, \boldsymbol{\Sigma}_{k+3}) = \mathcal{N}(\mu(\mathbf{x}_{k+3}^s), \sigma^2(\mathbf{x}_{k+3}^s))$



Izbira števila vzorcev  $S$  je odvisna od sistema, ki ga želimo opisati (red, kompleksnost odziva). Večje je število vzorcev  $S$ , bolj natančna, a hkrati bolj numerično potratna in s tem časovno daljša je simulacija.

Psevdo-koda opisanega postopka je v Dodatku C.1. Več o simulaciji GP modela s propagacijo negotovosti z numeričnimi MCMC metodami si lahko bralec prebere v [27], kjer lahko najdemo tudi primerjavo med rezultati numerične in analitične aproksimacije propagacije negotovosti.

### 3.5.4 Smotrnost propagacije negotovosti

Predhodno smo predstavili metode za simulacijo GP modelov z iterativno enokoračno predikcijo, ki se, kot smo že omenili, ločita na dve metodi:

- “naivna” metoda brez propagacije negotovosti in
- metoda s propagacijo negotovosti.

Kdaj uporabiti katero od metod?

Metoda s propagacijo negotovosti daje bolj natančno mero zaupanja v napoved, a je hkrati težja za izvedbo in računsko zahtevnejša tako za analitično (Dodatek B) kot za numerično propagacijo. Po drugi strani je metoda brez propagacije negotovosti hitra, enostavna in še vedno daje neko informativno mero zaupanja. Pas zaupanja, ki ga določa vrednost odmika od srednje vrednosti v velikosti dveh standardnih deviacij, pri “naivni” simulaciji sicer ne ustreza področju, v katerem lahko s 95% gotovostjo pričakujemo izhod sistema, ker je izračunano zaupanje preoptimistično. Kljub temu je širjenje in oženje področja zaupanja indikator zaupanja v model, ki ga lahko izkoristimo za analizo modela, podatkov ali pri raznih načrtovalnih postopkih.

Propagacija negotovosti ne daje samo bolj natančnega napovedovanja izhoda, marveč spremeni tudi srednjo vrednost napovedane porazdelitve. Odvisno od nelinearnosti sistema so te razlike lahko večje ali manjše. Ker gre tako pri simulaciji brez propagacije negotovosti in z njo za približek, lahko samo domnevamo, za koliko so rezultati slednje bolj točni.

“Naivno” metodo torej uporabimo, ko nas zanimata hitrost in enostavnost, medtem ko varianco propagiramo, ko nas zanima natančnejša mera zaupanja v napoved in ko čas trajanja simulacije ni omejen.

### 3.6 Vrednotenje

Vrednotenje je zelo pomemben del identifikacijskega postopka, s katerim preverimo ujemanje matematičnega modela in obravnavanega sistema oz. procesa [72]. Pregled vrednotenja in postopkov, ki se uporabljajo, je zbran npr. v [36,72].

Kvaliteto modela lahko merimo na različne načine [36]. *Verodostojnost* pove, ali se model sklada z našim predhodnim znanjem o procesu. Zanima nas odgovor na dve vprašanji:

- ali je model “videti” verodostojen in
- ali se model “obnaša” verodostojno?

GP model je videti verodostojen, če imajo hiperparametri razumne vrednosti. Tako za GP model s kovariančno funkcijo (2.4) vemo, da morajo biti vse vrednosti hiperparametrov  $\Theta$  večje od nič. Pri preverjanju verodostojnosti “obnašanja” nas lahko zanima npr. reakcija modela na tipične vhodne signale, kar lahko ovrednotimo z vizualnim ovrednotenjem odziva modela.

Ko nas zanima *veljavnost* modela, nas zanima ujemanje med izhodom modela in procesa ali med vhomom v proces in odzivom inverznega modela med simulacijo. Lahko jo ocenimo:

- kvantitativno, pri čemer si pomagamo s cenilkami, in
- kvalitativno, pri čemer npr. vizualno ovrednotimo razlike med izhodom modela in procesa in si lahko pomagamo tudi z napovedano varianco modela.

Običajne cenilke merijo samo razliko med odzivoma modela in procesa pri simulaciji ali enokoračni predikciji. Primera takih cenilk sta npr. povprečna kvadratna napaka (angl. *mean square error*, v disertaciji označena s SE) in povprečna relativna kvadratna napaka (angl. *mean relative square error*, MRSE) [100]. Pri GP modelu je smotrnejša uporaba cenilk, ki upoštevajo celotno (predpostavljeno) Gaussovo porazdelitev izhoda, saj se kot kriterij pri optimizaciji hiperparametrov upošteva največja verjetnost učnih podatkov (angl. *maximum likelihood*), ki prav tako upošteva celotno Gaussovo porazdelitev. Primer cenilke, ki upošteva celotno porazdelitev, je logaritem napovedane gostote napake (angl. *log-predictive density error*, LD) [27,47]; ta bolj uteži napako tistih napovedi, pri katerih model z manjšo varianco izkazuje večje zaupanje v napoved. Omenjene

cenilke so zbrane v naslednjih enačbah:

$$\text{SE} = \frac{1}{N} \sum_{i=1}^N e_i^2, \quad (3.20)$$

$$\text{MRSE} = \sqrt{\frac{\sum_{i=1}^N e_i^2}{\sum_{i=1}^N y_i^2}}, \quad (3.21)$$

$$\text{LD} = \frac{1}{2} \log(2\pi) + \frac{1}{2N} \sum_{i=1}^N \left( \log(\sigma_i^2) + \frac{e_i^2}{\sigma_i^2} \right), \quad (3.22)$$

kjer so  $y_i$ ,  $e_i = \hat{y}_i - y_i$  in  $\sigma_i^2$  izhod procesa, napaka modela in ocenjena varianca v  $i$ -tem koraku simulacije ali enokoračne predikcije.

Prvi pokazatelj kvalitete GP modela pa je že negativni logaritem verjetnosti učnih podatkov (angl. *negative log-likelihood of the training data*, LL) [27]:

$$\text{LL} = \frac{1}{2} \log |\mathbf{K}| + \frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} + \frac{N}{2} \log(2\pi), \quad (3.23)$$

kjer je  $\mathbf{K}$  kovariančna matrika GP modela,  $\mathbf{y}$  vektor učnih izhodov in  $N$  število učnih točk. Vrednost te cenilke pove, kakšna je “verjetnost”, da so podatki, uporabljeni za učenje GP modela, rezultat tistega naključnega procesa, ki ga naučeni GP model opisuje. Ta cenilka se uporablja v optimizaciji hiperparametrov, enačbi (2.10) in (2.11), in je na voljo takoj po končanem učenju modela. Za vse omenjene cenilke velja, da manjše kot so, boljši je ocenjevani GP model.

Tudi na samo varianco izhoda GP modela lahko gledamo kot na način vrednotenja, saj pove, kakšno je zaupanje v napoved na nekem področju obratovanja oz. kako dobro je model z učnimi podatki na tem področju opisan. Pri uporabi variance v vrednotenju moramo biti previdni, saj napovedi z majhno varianco niso nujno dobre, kot bo predstavljeno na primeru v podpoglavju 3.7.

Zadnji omenjeni vidik vrednotenja je *namembnost* (uporabnost) modela. Tu ocenjujemo, ali model ustreza namenu, zaradi katerega je bil zgrajen (definicija namena modela je prvi korak identifikacije), oz. odgovorimo na vprašanje, ali model lahko uporabljamo za namenjeno nalogo. Model je ovrednoten pozitivno, če z njim lahko rešimo nalogo, ki smo si jo pred modeliranjem zadali. Z vidika uporabnosti modela je to najvažnejši del vrednotenja. Tudi tu si lahko pomagamo z varianco, npr. model, namenjen napovedovanju odziva, katerega zaupanje v napovedane vrednosti je premajhno, lahko označimo za neuporaben, kot bo pokazano na primeru vrednotenja modela rasti alg v podpoglavju 3.8.

Glavna razlika v vrednotenju med GP modelom in bolj pogostimi modeli je uporaba variance [10, 12]. Ta je med vrednotenjem uporabna predvsem pri

vrednotenju veljavnosti in namembnosti GP modela. Kot smo že omenili, si lahko za ovrednotenje pomembnosti posameznih regresorjev pomagamo z ARD lastnostjo GP modela (podpoglavje 3.4), kar lahko koristno uporabimo pri izbiri regresorjev modela, kot bo pokazano v primerih v podpoglavjih 3.7 in 5.3.

### 3.7 Primer identifikacije modela bioreaktorja

Namen tega primera je demonstrirati postopek identifikacije GP modela, pri čemer bo posebna pozornost namenjena uporabi napovedane variance in drugih, za GP model značilnih, cenilk za vrednotenje modela. S primerom bomo najprej pokazali, kako so rezultati vrednotenja kriterij za izbiro modela. Na izbranem modelu bomo nato pokazali, kako na model vpliva povečan šum na izhodu iz sistema, kako se spremeni napovedovanje v nemodeliranih območjih in kako se model obnaša, če sistemu dodamo nemodeliran vhod.

Za predstavitev identifikacije si izberemo diskretizirani model bioreaktorja [16]. Število mikroorganizmov v bioreaktorju se povečuje z izrabo substrata. Bioreaktor je dan kot diskretni dinamični sistem drugega reda [16] s časom vzorčenja  $T_s = 0.5$  s:

$$\begin{aligned}x_1(k+1) &= x_1(k) + 0.5 \frac{x_1(k)x_2(k)}{x_1(k) + x_2(k)} - 0.5u(k)x_1(k), \\x_2(k+1) &= x_2(k) - 0.5 \frac{x_1(k)x_2(k)}{x_1(k) + x_2(k)} - 0.5u(k)x_2(k) + 0.05u(k), \\y(k) &= x_1(k) + \epsilon(k),\end{aligned}\quad (3.24)$$

kjer je  $x_1(k)$  koncentracija mikroorganizmov in  $x_2(k)$  koncentracija substrata. Zunanji vhod v sistem je izhodni pretok  $u(k)$ , omejen med 0 in 1. Izhod sistema  $y(k)$  je koncentracija mikroorganizmov, motena z belim Gaussovimi šumom  $\epsilon(k)$  s standardno deviacijo  $\sigma = 0.0005$ . Naša naloga je identifikacija tega sistema z GP modelom za napovedovanje obnašanja in ovrednotenje dobljenega modela.

Podatki za identifikacijo so bili dobljeni s simulacijo sistema (3.24), pri čemer so bile za vhodni signal  $u(k)$  uporabljene štiri sekunde dolge stopnice z naključnimi vrednostmi amplitud med 0 in 0.7. Zgornja meja amplitude vhodnega signala  $u(k)$  je bila izbrana tako, da je del delovnega območja sistema ostal neopisan.

Bioreaktor je bil identificiran z modelom četrtega, tretjega in drugega reda. Pred učenjem modela so bili signali normalizirani med vrednosti  $-1$  in  $1$ . Iz normaliziranih signalov sta bili sestavljeni 602 točki za učenje GP modela.  $i$ -ta učna točka v koraku  $k$  za sistem  $L$ -tega reda,  $L = 2, 3, 4$ , je sestavljena iz regresorjev:

$$\mathbf{x}_i = [y_n(k-1), \dots, y_n(k-L), u_n(k-1), \dots, u_n(k-L)] \quad (3.25)$$

in vrednosti izhoda  $y_i = y_n(k)$ , kjer sta  $u_n$  in  $y_n$  normalizirana vhodni in izhodni signal. Izbrana je bila Gaussova kovariančna funkcija (2.4), saj je bil predpostavljen stacionarni sistem z gladkim izhodom.

Vsi trije identificirani GP modeli so bili ovrednoteni s simulacijo, pri čemer smo podatke za vrednotenje dobili s simulacijo sistema (3.24) s podobnim, a ne enakim signalom  $u(k)$  kot za pridobitev podatkov za identifikacijo.

Med GP modeli smo izbirali z metodo izbire regresorjev na podlagi vrednotenja. Za vrednotenje GP modelov so bile uporabljene cenilke LL na podatkih za identifikacijo in MRSE ter LD na podatkih za vrednotenje. Rezultati so zbrani v tabeli 3.1.

Glede na vrednosti cenilk MRSE in LD, ki sta ovrednotili rezultate simulacije s podatki za vrednotenje, lahko sklepamo, da sistem (3.24) najbolje opisuje GP model drugega reda. Ta je boljši tudi s stališča principa Occamove britve (angl. *Occam's razor principle*) [91], ki pravi, da je med dvema približno enakovrednima rešitvama boljša preprostejša.

Hiperparametri  $w_{x^i}$ , kjer je  $x$  signal,  $x = u, y$ , in  $i$  ustrezen premik,  $i = 1, \dots, L$ , izražajo vpliv posameznih regresorjev. Iz velikosti hiperparametra  $w_{y^1}$  v tabeli 3.1 lahko razberemo, da je vpliv regresorja  $y(k-1)$  zelo majhen za vse potencialne modele, zato ga iz izbranega modela drugega reda odstranimo.

Ta postopek izbire regresorjev je privedel do GP modela, v tabeli 3.1 označenega z ♠, z regresorji:  $[y(k-2), u(k-1), u(k-2)]$ . Iz vrednosti cenilk LD in MRSE lahko sklepamo, da se tak model obnaša še bolje kot prvotni model.

Rezultati simulacije nad podatki za vrednotenje za izbrani model drugega reda so prikazani na sliki 3.6: prikazana sta izhod iz modela in nepošumljen izhod iz sistema. Lahko opazimo, da je večino časa vrednost napovedane standardne deviacije  $\sigma$  okoli  $5 \cdot 10^{-4}$ , kar ustreza velikosti šuma na izhodu sistema. Napovedana varianca se poveča v korakih, kjer se spremeni vhod  $u$ , zaradi česar model napoveduje na področju, opisanem z manj učnimi točkami (problem opisovanja dinamike sistema v neravnotežnih področjih, str. 22); zato se poveča člen  $k(\mathbf{x}) - \mathbf{k}(\mathbf{x})\mathbf{K}^{-1}\mathbf{k}(\mathbf{x})$  enačbe (2.15). Prav tako opazimo, da napaka napovedi modela vedno ostane znotraj pasu 95% negotovosti, definirane med  $\pm 2\sigma$ . Ti rezultati vrednotenja bodo služili za referenco, ko si bomo ogledali, kako različni pogoji vplivajo na identifikacijo in napovedovanje GP modela.

**Tabela 3.1:** Vrednosti cenilk in hiperparametrov za različne GP modele bioreaktorja

Model red	Pod. id. LL*	Pod. vred.		Hiperparametri <sup>b</sup>									šum <sup>‡</sup>	
		MRSE*	LD*	$w_{y^4}$	$w_{y^3}$	$w_{y^2}$	$w_{y^1}$	$w_{u^4}$	$w_{u^3}$	$w_{u^2}$	$w_{u^1}$	$v$	$\sigma$	
4	-1628	$8.5 \cdot 10^{-3}$	-6.11	$1.2 \cdot 10^{-2}$	$2.6 \cdot 10^{-2}$	$4.3 \cdot 10^{-2}$	$2 \cdot 10^{-6}$	$6 \cdot 10^{-6}$	$1.7 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	$1.8 \cdot 10^{-2}$	6.0	$5.0 \cdot 10^{-4}$	
3	-1621	$7.5 \cdot 10^{-3}$	-6.12	×	$2.5 \cdot 10^{-3}$	$6.6 \cdot 10^{-2}$	$4 \cdot 10^{-4}$	×	$3.6 \cdot 10^{-3}$	$3.6 \cdot 10^{-3}$	$3.6 \cdot 10^{-3}$	18	$5.2 \cdot 10^{-4}$	
2	-1612	$6.1 \cdot 10^{-3}$	-6.41	×	×	$3.0 \cdot 10^{-2}$	$1.1 \cdot 10^{-3}$	×	×	$5.2 \cdot 10^{-3}$	$5.1 \cdot 10^{-3}$	28	$5.3 \cdot 10^{-4}$	
2♠	-1612	$4.4 \cdot 10^{-3}$	-6.42	×	×	$3.0 \cdot 10^{-2}$	×	×	×	$5.2 \cdot 10^{-3}$	$5.1 \cdot 10^{-3}$	28	$5.3 \cdot 10^{-4}$	
2♣	-793	$5.0 \cdot 10^{-3}$	-5.19	×	×	$4.3 \cdot 10^{-2}$	×	×	×	$9.2 \cdot 10^{-3}$	$9.2 \cdot 10^{-3}$	10.4	$2.1 \cdot 10^{-3}$	

Opombe:

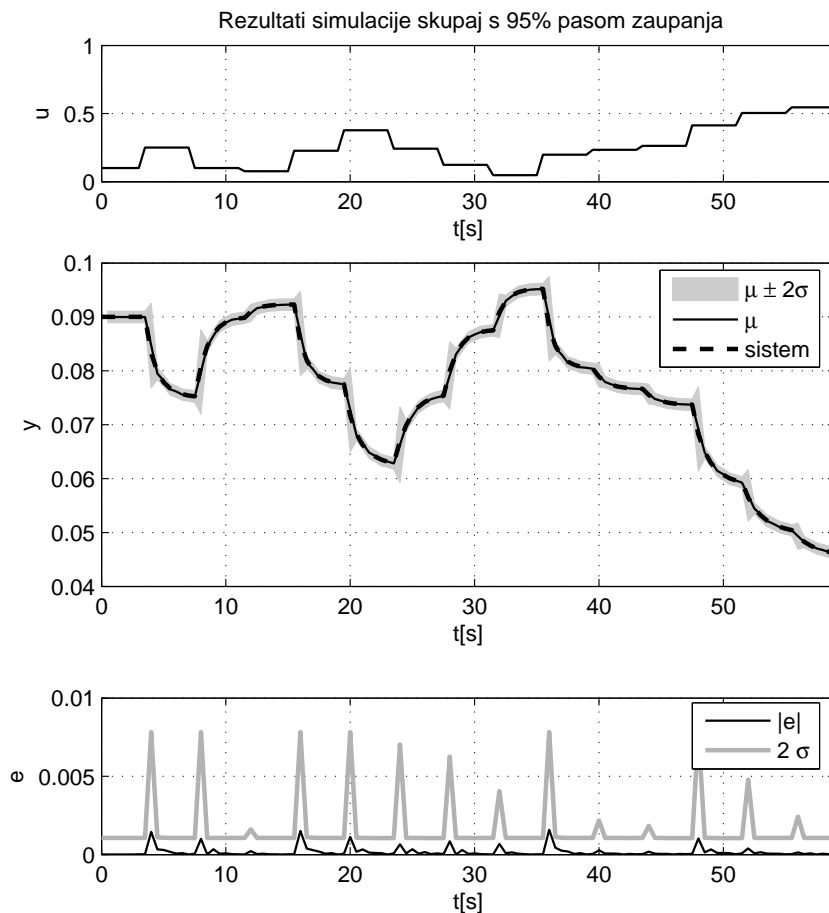
<sup>b</sup>  $w_{x^i}$  označuje hiperparameter, ki ustreza vzorcu  $x(k-i)$  signala  $x$

<sup>‡</sup> identificirana standardna deviacija šuma  $\sigma$

♠ model z odstranjenim regresorjem  $y(k-1)$

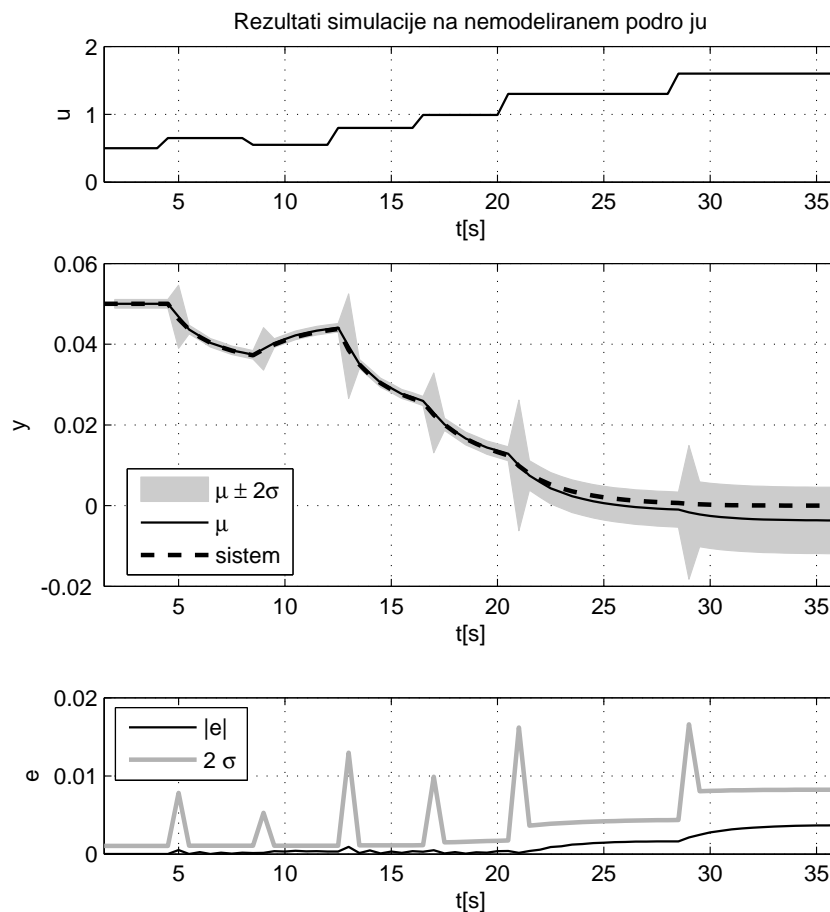
♣ model identificiran s podatki, vzorčenimi iz izhodnega signala z večjo varianco šuma  $\sigma = 0.002$

Najprej si bomo ogledali, kako se napoved modela spremeni, ko je model na nemodeliranem področju. Ker v bližini ni učnih podatkov, mora model ekstrapolirati napoved iz podatkov, ki opisujejo sosednja področja. To poslabša srednjo vrednost napovedi, a se izrazi tudi v povečanju napovedane variance oz. razširitvi pasa zaupanja. Ta efekt je prikazan na sliki 3.7, ki prikazuje odziv sistema in modela na povečane vrednosti vhoda  $u(k)$  nad 0.7 v času  $t > 12$  s.



**Slika 3.6:** Vrednotenje GP modela bioreaktorja s simulacijo

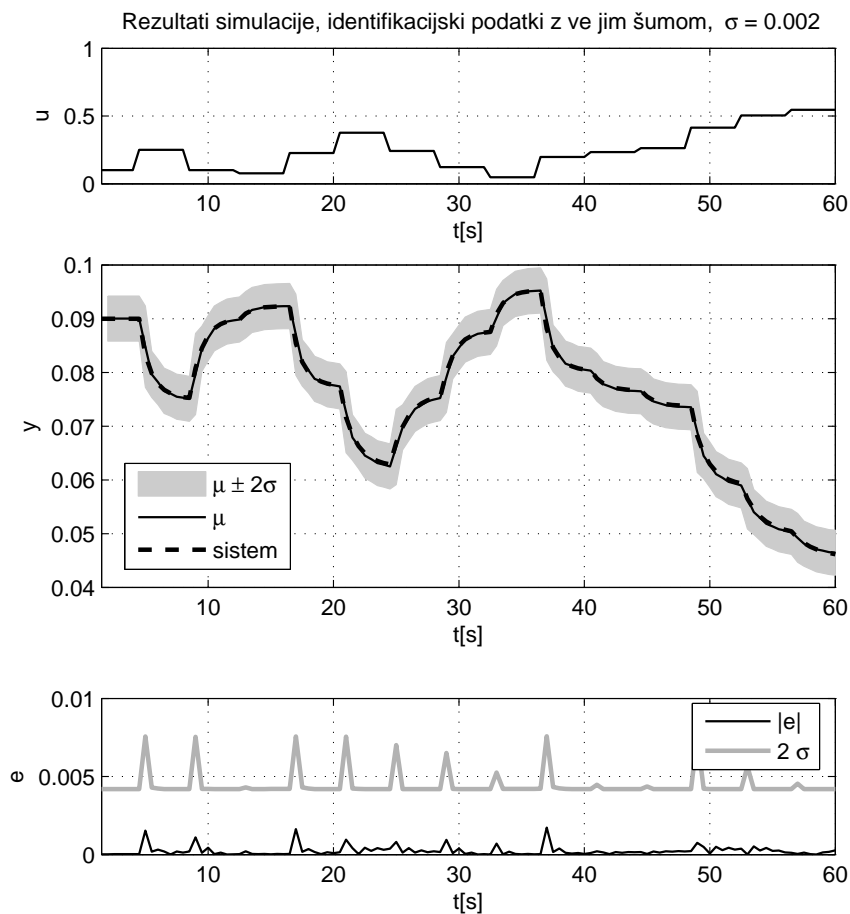
Drugič, ogledali si bomo, kako povečanje variance šuma na izhodu sistema vpliva na identificiran sistem. Za ta primer je bila vrednost standardne deviacije šuma izhoda povečana na  $\sigma = 2 \cdot 10^{-3}$ . Vhodna signala  $u$ , uporabljena za pridobitev podatkov za identifikacijo in vrednotenje, sta ostala ista kot v referenčnem primeru. Z novimi podatki je bil identificiran model drugega reda, iz katerega je bil prav tako odstranjen regresor  $y(k-1)$ . Vrednosti hiperparametrov tega GP modela so v tabeli 3.1 v vrstici, označeni z ♣.



**Slika 3.7:** Napovedovanje z uporabo GP modela na nemodeliranem področju

Srednja vrednost napovedi tega modela je zadovoljiva. Ocena velikosti šuma na izhodu sistema je zadovoljivo blizu prave vrednosti, tj.  $\sigma = 2.1 \cdot 10^{-3}$ , in napovedana varianca se ustrezno poveča na njen račun, kot je prikazano na sliki 3.8. Vrednost cenilke MRSE je nekoliko slabša kot za referenčni primer, saj je model identificiran z bolj pošumljenimi učnimi podatki. Prav tako je nekoliko slabša vrednost cenilke LD, saj kljub povečani vrednosti variance prevlada povečana vrednost napake simulacije.

Na koncu si bomo ogledali še, kako se GP model obnaša, če nanj vpliva nemodeliran vhod. Za ta namen je bil sistemu (3.24) dodan dodatni vhod  $z$ , nekoreliran z vhomom  $u$ . Vpliv tega vhoda je sicer enak vplivu vhoda  $u$  in bi



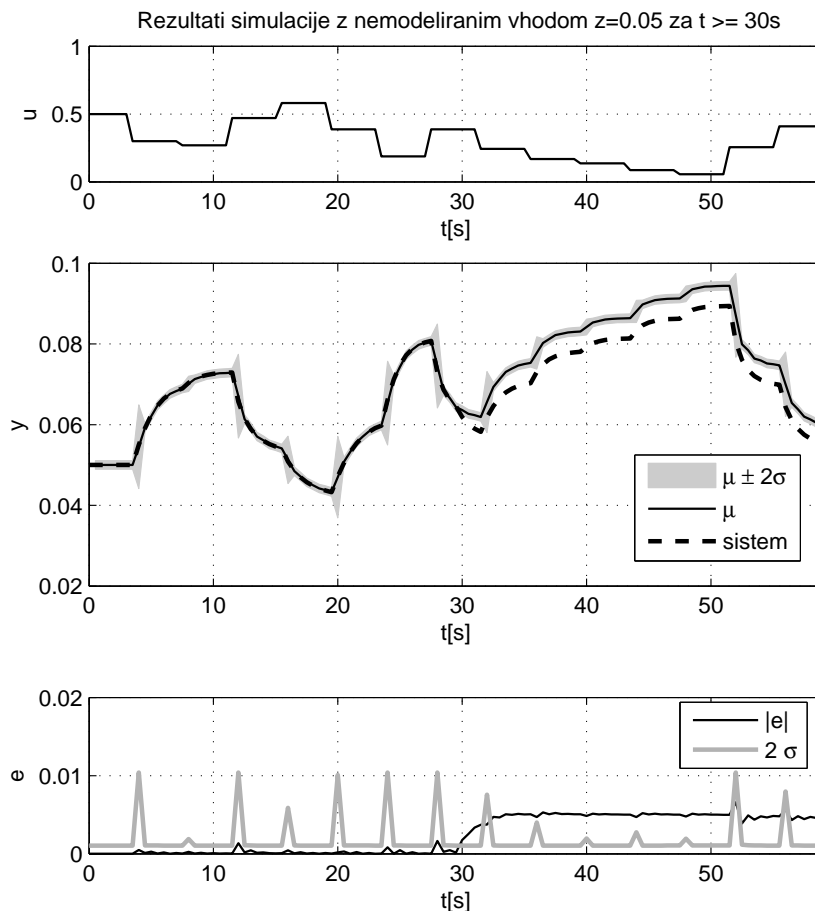
**Slika 3.8:** Vpliv povečanega izhodnega šuma sistema na napovedovanje

lahko predstavlja dodatni odtok ali puščanje sistema. Tak sistem opišemo z:

$$\begin{aligned}
 x_1(k+1) &= x_1(k) + 0.5 \frac{x_1(k)x_2(k)}{x_1(k) + x_2(k)} - 0.5u(k)x_1(k) - 0.5z(k)x_1(k), \\
 x_2(k+1) &= x_2(k) - 0.5 \frac{x_1(k)x_2(k)}{x_1(k) + x_2(k)} - 0.5u(k)x_2(k) - \\
 &\quad - 0.5z(k)x_2(k) + 0.05u(k) + 0.05z(k), \\
 y(k) &= x_1(k) + \epsilon(k).
 \end{aligned} \tag{3.26}$$

Za napovedovanje uporabimo referenčni GP model, pri katerem ob učenju vhoda  $z$  ni bilo in je bil pri učenju zanemarljiv. Med vrednotenjem modela s simulacijo vhod  $z$  sistema (3.26) spremeni vrednost iz nič na  $z = 0.05$  v času 30 s. (Ne)vpliv v GP modelu nezajetega vhoda  $z$  na napovedano varianco na sicer dobro opisanem področju je prikazan na sliki 3.9. Napoved modela se od časa 30 s naprej poslabša,

a 95% pas zaupanja ostane ozek. Ta primer kaže, da varianca ne more dati informacije o vplivih na sistem, ki med učenjem GP modela niso zajeti.



**Slika 3.9:** Vpliv nemodeliranega vhoda na napovedovanje

Primer identifikacije in vrednotenja bioreaktorja je ilustriral naslednje lastnosti GP modela:

1. ARD lastnost hiperparametrov GP modela lahko učinkovito uporabimo za zmanjšanje števila regresorjev identificiranega modela.
2. Napovedana varianca se lahko poveča zaradi dveh razlogov:
  - delovno območje, na katerem model napoveduje, je nezadostno opisano z učenimi podatki in
  - podatki, ki opisujejo delovno območje, so bolj pošumljeni. V primeru je bilo to pokazano za celotno delovno območje, vendar enako velja za primer, da se šum poveča le na določenem področju delovanja modela.

Razloga za povečano varianco sta težko ločljiva brez predhodnega znanja o sistemu.

3. Če je identificiran sistem podvržen nemodeliranemu vplivu, se napovedovanje modela ne more spremeniti.

### 3.8 Primer vrednotenja – Beneška laguna

Namen tega primera je pokazati, kako je za vrednotenje GP modela mogoče uporabiti vrednosti napovedane variance [10, 12].

Radi bi zgradili vhodno/izhodni model za napovedovanje koncentracije biomase v Beneškem zalivu. Zaliv meri v površino 550 km<sup>2</sup>, a je zelo plitek, s povprečno globino manj kot en meter. Na njegovo onesnaženost močno vpliva dotok nitratov in fosfatov človeškega in industrijskega izvora: približno sedem milijonov kg/leto dušika in 1.4 milijone kg/leto fosforja [109]. Ta dotok (večinoma) nitratov povzroča prekomerno rast alg, pretežno vrste *Ulva rigida*.

Za modeliranje rasti alg v zalivu so bile na voljo štiri razmeroma majhne množice meritev [17]. Meritve so se opravljale enkrat tedensko za obdobje malo daljše od enega leta na štirih različnih lokacijah v zalivu in so močno pošumljene z meritveno napako od 20 do 50% [109]. Merjena je bila koncentracija dušika v amoniju NH<sub>3</sub>, dušika v nitratih NO<sub>3</sub>, skupna koncentracija dušika N, koncentracija fosforja v ortofosfatih PO<sub>4</sub>, vse v [μg/l], koncentracija raztopljenega kisika DO v procentih nasičenja, temperatura T v [°C] in biomasa alg B, čista teža v [g/m<sup>2</sup>].

Namen modela je napovedovanje rasti alg, zaradi česar je izhod iz modela biomasa alg B. Preostale merjene spremenljivke so potencialni vhodi v model. Ker je število učnih podatkov zelo majhno, je bil predpostavljen sistem prvega reda.

Namen primera je pokazati uporabnost variance za vrednotenje modela, zato sta bili za identifikacijo uporabljeni samo dve množici 43 meritev, zbranih ob istem času, a na različnih lokacijah – ena za učenje GP modela in druga za njegovo vrednotenje. Pred uporabo podatkov so bile meritve normalizirane.

Za izbiro regresorjev modela je bila uporabljena metoda izbire regresorjev na podlagi vrednotenja. Preizkušenih je bilo več modelov prvega reda z različnimi kombinacijami regresorjev. Rezultatih enokoračne predikcije na podatkih za vrednotenje so bili ovrednoteni vizualno in z uporabo cenilk MRSE (3.21) in LD (3.22). Ker učni podatki nosijo malo informacije in ker so bile razlike med identificiranimi modeli majhne, je bila izbira regresorjev otežena.

Z izbiro regresorjev na podlagi vrednotenja je bil za napovedovanje biomase  $B$  v koraku  $k$  izbran GP model z naslednjimi regresorji:

$$\mathbf{x}_k = [B(k-1) \ T(k-1) \ DO(k-1) \ N(k-1)] \quad (3.27)$$

z vrednostmi hiperparametrov:  $w_1 = 177$ ,  $w_2 = 6.6 \cdot 10^{-7}$ ,  $w_3 = 32$ ,  $w_4 = 5$ ,  $v = 2.8$  in  $v_0 = 1.0 \cdot 10^{-3}$ , kjer vrednost  $v_0$  ustreza standardni deviaciji belega šuma  $\sigma = 33.8$ . Zaradi majhne vrednosti ustreznega hiperparametra je bil iz modela odstranjen regresor  $T(k-1)$ , s čimer smo dobili model z regresorji:

$$\mathbf{x}_k = [B(k-1) \ DO(k-1) \ N(k-1)]. \quad (3.28)$$

Rezultati enokoračne predikcije na podatkih za identifikacijo in podatkih za vrednotenje so prikazani na sliki 3.10, rezultati simulacije na podatkih za vrednotenje pa na sliki 3.11.

Model dobro opisuje podatke za identifikacijo, vendar rezultati enokoračne predikcije in simulacije na podatkih za vrednotenje izkazujejo visoko vrednost cenilke MRSE, a so primerljivi z rezultati drugih poizkusov modeliranja istega sistema, npr. [109]<sup>4</sup>, kjer so rezultati ocenjeni kot dobri, če so vrhovi in doline koncentracije biomase napovedani pravilno.

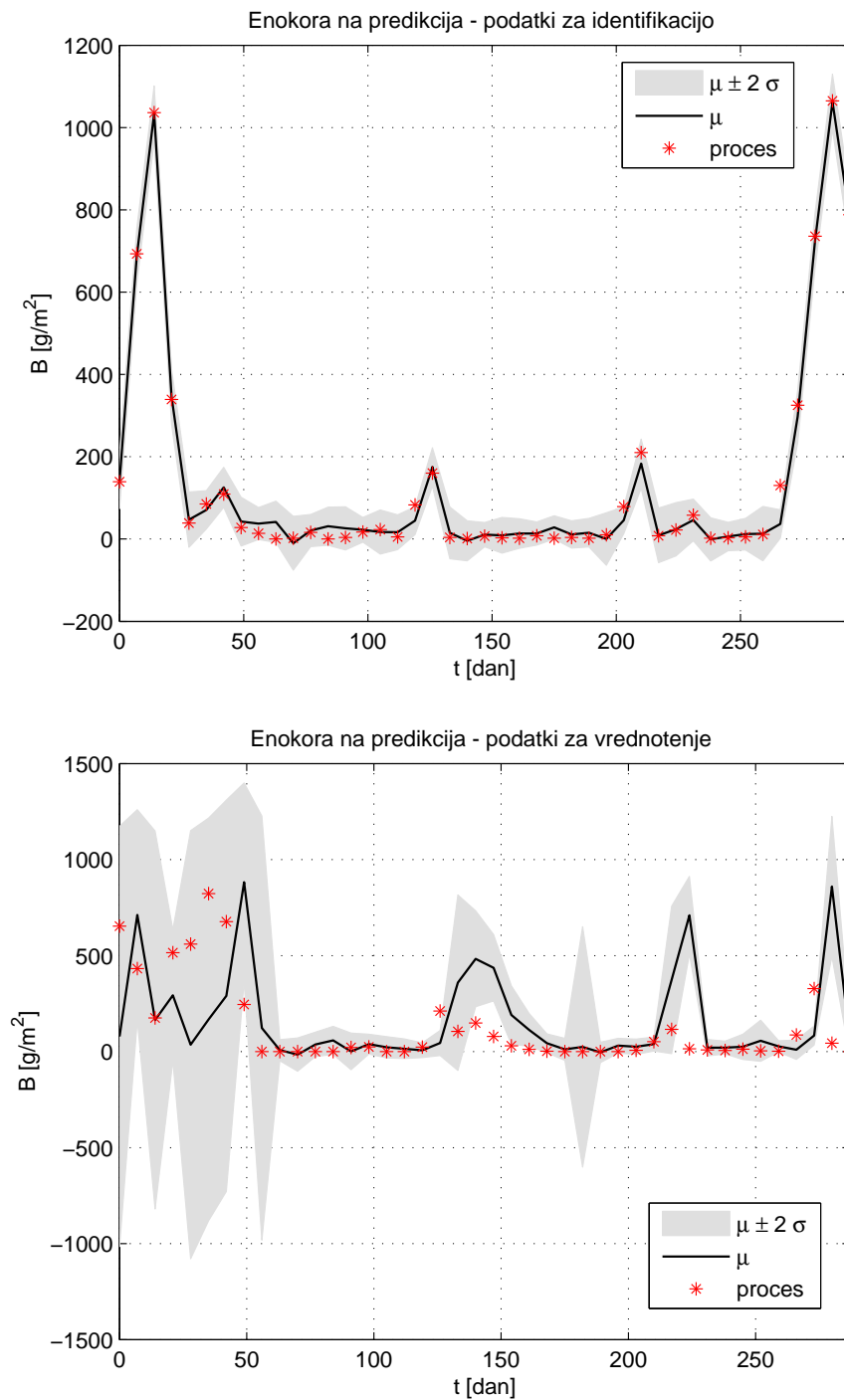
Če pa si ogledamo 95% pas zaupanja na slikah 3.10 (spodaj) in 3.11, ki opisujeta rezultate enokoračne predikcije in simulacije nad podatki za vrednotenje, vidimo, da je zaupanje modela v napovedi zelo majhno<sup>5</sup>. Na podlagi vrednosti napovedanih varianc se odločimo, da modelu ne moremo zaupati, kljub temu da so napovedane srednje vrednosti potencialno sprejemljive. Razpoložljiva mera zaupanja se izkaže kot priročno orodje za vrednotenje.

Vrednosti cenilk za prikazane rezultate so zbrane v tabeli 3.2. Iz rezultatov vidimo, da se pri vrednotenju s podatki, ki niso bili uporabljeni za identifikacijo, model obnaša bolje pri simulaciji kot pri enokoračni predikciji, kar nakazuje sumljivo kvaliteto identificiranega modela.

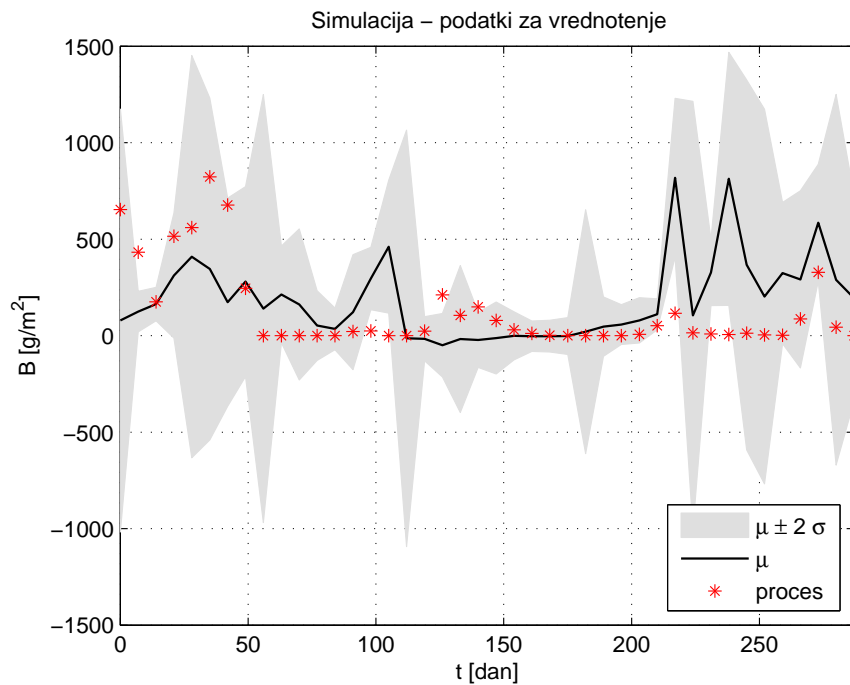
Tudi iz razmeroma nizke napovedane ocene standardne deviacije  $\sigma = 33.8$  lahko sklepamo, da bi za učenje morali uporabiti več podatkov, saj napovedana varianca ne ustreza omenjeni oceni napak meritev, kot npr. v primeru bioreaktorja, podpoglavje 3.7, kjer je povečanje šuma na izhodu sistema spremljala povečana varianca napovedi GP modela.

<sup>4</sup>V [109] je bila uporabljena kombinacija pristopa k modeliranju na podlagi znanja (angl. *knowledge-driven approach*) in podatkov (angl. *data-driven approach*). Za opis stanja  $B$  sta bila v enem primeru uporabljena regresorja  $DO$  in  $NH_3$ , v drugem pa še dodatni regresor  $T$ .

<sup>5</sup>Z upoštevanjem predhodnega znanja bi lahko omejili napovedovanje in zaupanje modela samo nad fizično možno vrednost  $B \geq 0$  g/m<sup>2</sup>, vendar bi to zmanjšalo ilustrativnost primera.



**Slika 3.10:** Primerjava enokoračne predikcije nad podatki za identifikacijo (zgoraj) in podatki za vrednotenje (spodaj)



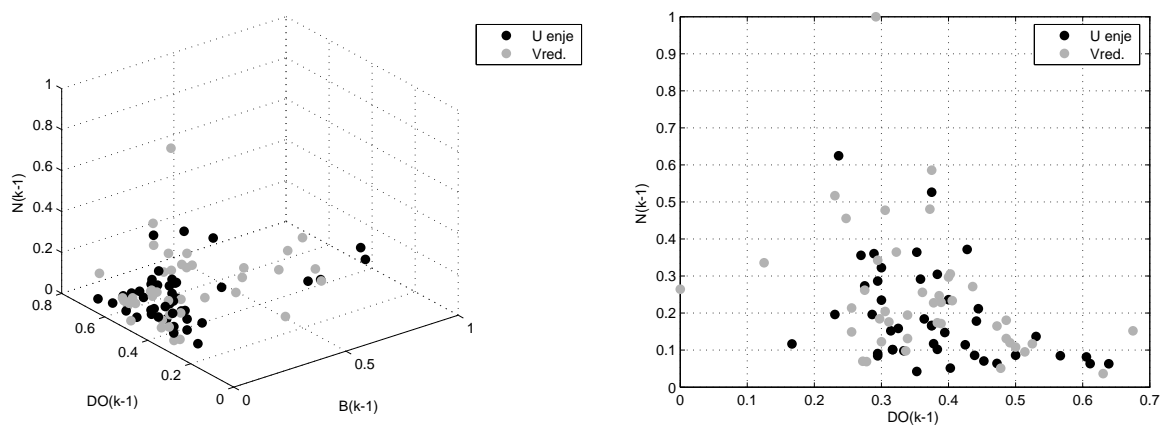
**Slika 3.11:** Simulacija GP modela nad podatki za vrednotenje

O pomanjkanju učnih podatkov na določenih področjih lahko sklepamo tudi iz slike 3.12, iz katere razberemo, da razmeroma malo učnih podatkov opisuje proces pri večjih koncentracijah biomase  $B$ . Zato je napovedovanje rasti biomase na teh področjih oteženo, kar je razvidno iz slik 3.10 (spodaj) in 3.11, vendar je manj natančno napovedovanje spremljano s povečano varianco na tem območju.

Ta primer je predstavil uporabo variance za vrednotenje dobljenega GP modela. Model bi lahko izboljšali z uporabo večjega števila podatkov za identifikacijo, s čimer pa bi zmanjšali nazornost primera.

**Tabela 3.2:** Vrednosti cenilk za rezultate enokoračne predikcije in simulacije

	Podatki za ident. enokoračna	Podatki za vrednotenje enokoračna	simulacija
MRSE	0.068	1.13	1.10
LD	4.7	8.3	7.4



Slika 3.12: Primerjava lege podatkov za vrednotenje in identifikacijo v treh dimenzijah (levo) in dveh dimenzijah, ko vrednosti regresorja  $B(k-1)$  niso prikazane (desno)

## Poglavje 4

# Vključevanje predznanja v GP model

V prejšnjem poglavju smo predstavili identifikacijo nelinearnih dinamičnih sistemov z GP modelom. Pri tem smo predpostavili, da o identificiranem sistemu nimamo predznanja ali pa je tega zelo malo. Če neko predznanje o sistemu obstaja, je smiselno in skladno z osnovnimi principi identifikacije<sup>1</sup>, da poskusimo to znanje vključiti v model.

Predznanje o sistemu, ki ga želimo modelirati, upoštevamo že pri izbiri kovariančne funkcije GP modela in regresorjev sistema, ki nastopajo kot vhodi v GP model. V tem poglavju bomo govorili o eksplicitnem vključevanju posebnih oblik predznanja, značilnih za dinamične sisteme, v GP model [5].

Najprej bosta predstavljena dva načina vključevanja predznanja: (a) spreminjanje kovariančne funkcije in (b) spreminjanje vhodno/izhodnih podatkov. Nato bosta predstavljeni dve pomembnejši različici GP modelov, ki jih dobimo z vključevanjem predznanja. Prvi tak model je GP model z vključenimi lokalnimi modeli, ki nastane z vključevanjem lokalnih modelov v GP model. Drugi tak model je GP model z nespremenljivo strukturo. Temu modelu strukturo določimo eksplicitno, medtem ko z GP modeli modeliramo samo njegove parametre. V preostanku poglavja bo sledila še predstavitev vključevanja predznanja o histerezi in predznanja o šumnem procesu na izhodu sistema.

Vključevanje predznanja lahko pomembno vpliva na lastnosti GP modela. Navadno izboljša delovanje modela, saj model zaradi vključenega predznanja lahko daje bolj točne napovedi, npr. pri vključevanju predznanja o histerezi ali dinamičnem modelu šuma. Z vključevanjem znanja se lahko pohitri optimizacija in napovedovanje modela zaradi zmanjšanja učne množice ali poveča

---

<sup>1</sup>“*Ne identificiraj tistega, kar že znaš.*”

transparentnost neparametričnega GP modela, npr. pri vključevanju lokalnih modelov itd. Predstavili bomo vključevanje:

- statične karakteristike [4, 46],
- linearnih lokalnih modelov [4, 46, 99],
- histereze [8],
- znanja o šumnem procesu na izhodu [11, 73] in

si pogledali, kako lahko uporabimo znanje o strukturi opisovanega sistema za identifikacijo parametričnega modela, temelječega na GP modelih [9].

Za lažjo primerjavo med GP modeli različnih postavitev bomo model, kot je bil predstavljen do sedaj, tj. brez eksplicitno vključenega predznanja in z Gaussovo kovariančno funkcijo, imenovali “običajni” GP model.

## 4.1 Načini vključevanje predznanja

Informacija o identificiranem sistemu je v GP modelu vsebovana v dveh oblikah:

- v podatkih  $\mathcal{D} = \mathbf{X}, \mathbf{y}$ , ki opisujejo vhodno/izhodno obnašanje sistema in
- v kovariančni funkciji oz. kovariančni matriki, s katero je izražena odvisnost med podatki.

Tako se ponujata dve možnosti za vključevanje predznanja v GP model. Prva možnost je taka sprememba kovariančne funkcije (oz. matrike), da ta ustreza predznanju o sistemu. Spremenimo lahko družino kovariančne funkcije, vrednosti hiperparametrov ali pa kar direktno celo ali del kovariančne matrike. Druga možnost vključevanja predznanja je dodajanje ali sprememba vhodno/izhodnih učnih vektorjev zbranih v  $\mathcal{D}$ .

### 4.1.1 Spreminjanje kovariančne funkcije/matrike

Naloga kovariančne funkcije v GP modelu je določitev odvisnosti med podatki, ki sestavljajo GP model. Z izbiro družine kovariančne funkcije izrazimo naše predznanje. Družina funkcij (2.4) je široko uporabljena, saj izraža mnogokrat predpostavljene lastnosti, kot so stacionarnost opisovanega procesa in gladkost

njegovega izhoda, in je poleg tega še relativno lahka za uporabo. Prva možnost vključevanja predznanja je izbira druge družine kovariančnih funkcij, prim. [26, 91], ko vemo, da opisujemo proces z drugačnimi lastnostmi, kot so nestacionarnost, periodičnost [58] itd.

Naslednja možnost je sprememba narave nekaterih podatkov v učnih podatkih  $\mathcal{D}$ . Pri tem je za nov tip podatkov potrebno spremeniti tudi kovariančno funkcijo. Če podatki nosijo informacijo o odvodih in ne o vrednostih funkcije, lahko v GP model vključimo linearne lokalne modele. Tak model imenujemo GP model z vključenimi lokalnimi modeli (LMGP model), predstavljen bo v podpoglavju 4.2.

Mnogokrat se zgodi, da je izhod moten z barvnim in ne (s pogosto predpostavljenim) belim šumom. V tem primeru lahko spremenimo “šumni” del kovariančne matrike  $\mathbf{K}$ , kot bomo pokazali v razdelku 4.4.3. Druge možnosti modeliranja šuma so povzete v [91] in v razdelku o izbiri kovariančne funkcije na strani 35.

#### 4.1.2 Spreminjanje vhodno/izhodnih podatkov

Poleg spreminjanja kovariančne funkcije/matrike, lahko predznanje vključujemo tudi z dodajanjem ali spreminjanjem vhodno/izhodnih podatkov, zbranih v  $\mathcal{D} = \mathbf{X}, \mathbf{y}$ , s katerimi je obnašanje neznanega procesa izraženo v eksplicitni obliki, tj. odziv procesa kot funkcija ustrezajočih vrednosti vhodnih regresorjev. Ponuja se več različnih možnosti:

- Prva možnost je dodajanje novih učnih točk  $\mathbf{x}_i|y_i$  že obstoječim, s čimer lahko v GP model vključimo predhodno znanje npr. o statični karakteristiki ali eksplicitnih omejitvah procesa. Tako bi npr. za dinamični model rasti neke populacije v odvisnosti od hrane lahko eksplicitno vključili pogoj: če v sistemu za dalj časa zmanjka hrane, vsa od nje odvisna populacija pogine. Drugi primer bi lahko bila omejitev minimalne koncentracije neke snovi v procesu na nič, saj negativne koncentracije niso mogoče.

Problem, ki se lahko pojavi, je, kako tako predznanje vključiti v model z večjim številom regresorjev  $D$ . Vsaka vključena točka namreč predstavlja eno točko v vhodnem  $D$ -dimenzionalnem prostoru in za dobro predstavo še vedno potrebujemo precejšnje število točk.

Primer vključevanja predznanja o statični karakteristiki bo prikazan skupaj z vključevanjem linearnih lokalnih modelov v podpoglavju 4.2.

- Druga možnost je, da že uporabljenim regresorjem dodamo nove, ki jih uporabimo za opis dodatnega znanja o procesu, s čimer povečamo vhodno

dimenzijo. Primer informacije, ki jo nov regresor lahko nosi, je npr. stanje histereze, katere vključevanje bomo na primeru predstavili v razdelku 4.4.2.

- Zadnja v tej disertaciji predstavljena možnost se precej razlikuje od opisanih. Če poznamo strukturo opisovanega procesa, tj. njegove regresorje in red  $L$ , lahko za opis procesa uporabimo linearni model s spremenljivimi parametri (angl. *linear-parameter-varying model*, LPV model, model z linearno strukturo) [14, 94]. Ko smo za napovedovanje vrednosti spremenljivih parametrov uporabili GP modele, smo tak model imenovali GP model z vnaprej določeno strukturo (angl. *Fixed-Structure GP model*, FSGP model). Ta pristop, podoben pristopu z LMN, bo predstavljen v pod poglavju 4.3.

\* \* \*

V naslednjih pod poglavjih bodo nekateri prej naštetih načini vključevanja predznanja v GP model predstavljeni na primerih dveh nelinearnih dinamičnih sistemov. Prvi sistem je bil uporabljen za predstavitev delovanja ANN za identifikacijo nelinearnih dinamičnih sistemov [78] in je zato primeren za primerjavo. Gre za nelinearni dinamični sistem prvega reda, opisan z enačbo:

$$y(k+1) = \frac{y(k)}{1+y^2(k)} + u^3(k). \quad (4.1)$$

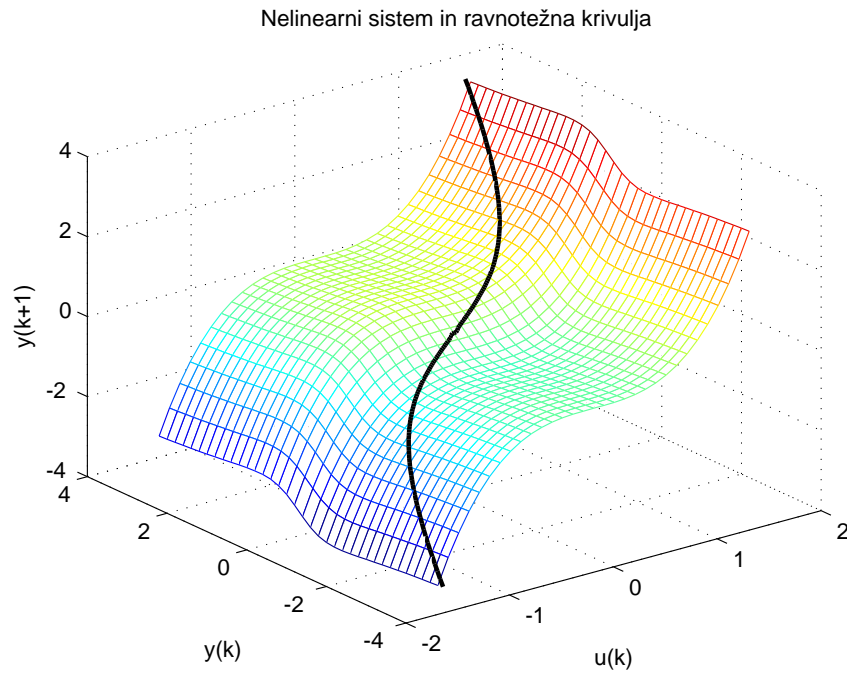
Odziv sistema je za vhode omejene med  $u_{\min} = -1.5$  in  $u_{\max} = 1.5$  prikazan na sliki 4.1. S črno črto je označena njegova ravnotežna krivulja.

Z drugim primerom bomo ilustrirali identifikacijo GP modela z vključenimi lokalnimi modeli in GP modela z vnaprej določeno strukturo. Uporabili bomo nelinearni dinamični sistem drugega reda [71]:

$$y(k) = 0.893 y(k-1) + 0.0371 y^2(k-1) - 0.05 y(k-2) - 0.05 u(k-1) y(k-1) + 0.157 u(k-1) + \epsilon(k), \quad (4.2)$$

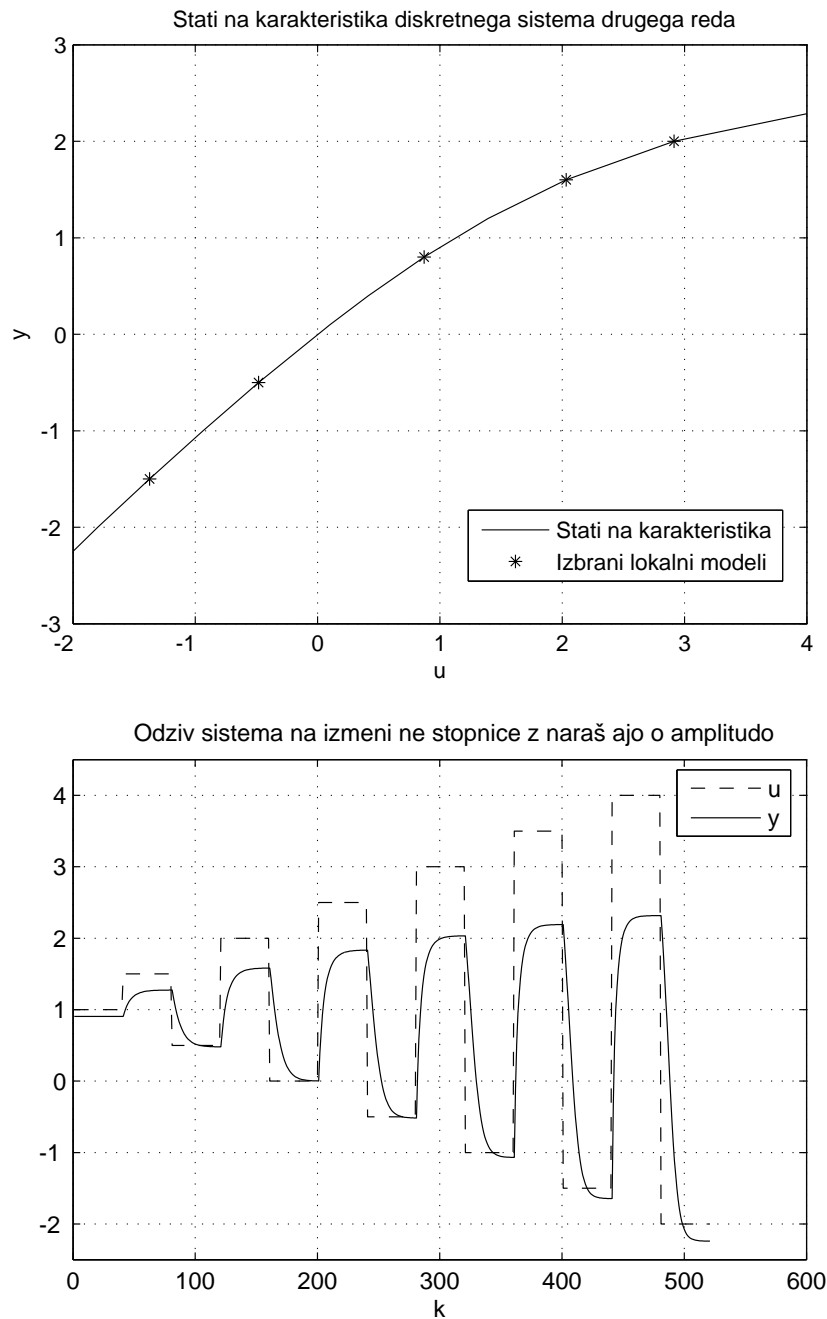
kjer je  $\epsilon(k)$  beli Gaussov šum z varianco  $\sigma^2$ . Sistem ima tri regresorje, zato ga skupaj z izhodom ne moremo predstaviti v tridimenzionalnem prostoru, kot smo predstavili sistem prvega reda (4.1).

Statično karakteristiko sistema (4.2) na področju, omejenem z vhomom  $u$  med  $u_{\min} = -2$  in  $u_{\max} = 4$ , si lahko ogledamo na sliki 4.2 zgoraj. Na sliki 4.2 spodaj je prikazan odziv sistema na izmenične stopnice z naraščajočo amplitudo okoli



**Slika 4.1:** Odziv nelinearnega dinamičnega sistema prvega reda (4.1) in njegova ravnotežna krivulja

vrednosti  $u = 1$ , na katerem vidimo, da se sistemu spreminja dinamika časovnega odziva glede na vrednosti vhodnega signala.



**Slika 4.2:** Statična karakteristika (zgoraj) in odziv nelinearnega dinamičnega sistema drugega reda (4.2) na izmenične stopnice z naraščajočo amplitudo (spodaj), namenjena ilustraciji delovanja LMGP in FSGP modela

## 4.2 GP model z vključenimi lokalnimi modeli

V poglavju 3.1 smo govorili o različnih pristopih k identifikaciji nelinearnih dinamičnih sistemov. Pri tem smo omenili tudi mrežo lokalnih modelov (LMN), pri kateri je neznan sistem opisan z množico lokalnih modelov, ki jih glede na stanje sistema zlivamo in tako opišemo sistem v celotnem identificiranem področju. Izpostavili smo naslednje probleme tega pristopa:

1. problem opisovanja dinamike v neravnotežnih področjih;
2. izbira med opisom globalne ali lokalne dinamike z LMN, pri čemer je problem prvega pristopa neverodostojen opis dinamike z lokalnimi modeli na neravnotežnem področju, problem drugega pa je majhnost področja veljavnosti posameznih lokalnih modelov;
3. problem izbire vektorja razvrstilnih spremenljivk;
4. problem razdelitve območja delovanja sistema med lokalne modele.

Kot smo omenili, je ena izmed možnosti za zmanjšanje nekaterih od teh problemov uporaba VBL pristopa (podpoglavje 3.1), namenjenega zveznim modelom. Vendar tudi ta za opis nelinearnega sistema potrebuje lokalne modele na področjih, kjer je podatkov za njihovo gradnjo malo. To je bil eden izmed razlogov za iskanje alternativnih metod identifikacije nelinearnih dinamičnih sistemov in predlagana je bila uporaba GP modela [56].

Kljub vsem problemom, ki nastanejo pri modeliranju z LMN, je uporaba tega pristopa zaradi navidezne transparentnosti še vedno vabljiva. Zato je bil za identifikacijo nelinearnih dinamičnih sistemov predlagan GP model, v katerega bi bili vključeni lokalni modeli [56,99]. Tak GP model bomo imenovali *GP model z vključenimi lokalnimi modeli* (angl. *GP model with incorporated Local Models*, LMGP model) [4,6].

V [56] je bila predlagana uporaba LMGP modela za opisovanje dinamičnih sistemov. Izvedba vključevanja lokalnih modelov za Gaussovo kovariančno funkcijo (2.4) je bila predstavljena in ilustrirana s primerom v [99]. Ilustracija skupaj z uporabo LMGP modela za vodenje je bila predstavljena v [49]. V [13] je bila ilustrirana uporaba LMGP modela za prediktivno vodenje z omejitvami, pri katerih si pomagamo z varianco napovedi modela. Identifikacija realne naprave z LMGP modelom je predstavljena v [4,44]. Propagacijo negotovosti skozi LMGP model z analitično aproksimacijo najdemo v [46], propagacijo z numerično aproksimacijo pa v [6].

### 4.2.1 Linearni lokalni modeli

Linearni lokalni model opisuje lokalno obnašanje sistema, tj. obnašanje sistema na nekem podpodročju, z linearnim modelom. V primeru opisovanja nelinearnega sistema je pogosta praksa, da razbijemo področje delovanja na podpodročja, na katerih predpostavljamo linearno obnašanje sistema in zato za opis vsakega podpodročja uporabimo (linearni) lokalni model. Lokalne modele nato združimo v globalen model, ki opisuje sistem na celotnem območju delovanja. To je lahko npr. LMN, mehki model ali, kot bo prikazano v tem podpoglavju, LMGP model.

Zanima nas diskretni opis nelinearnega dinamičnega sistema:

$$y(k) = f(\mathbf{x}_k) \quad (4.3)$$

$$\mathbf{x}_k = [y(k-1) \dots y(k-L) u(k-1) \dots u(k-L)]^T. \quad (4.4)$$

Predpostavimo, da je točka (tj. vektor regresorjev oblike (4.4))  $\mathbf{x}_m$  ravnotežna točka tega sistema. Radi bi opisali dinamiko sistema v bližini ravnotežne točke  $\mathbf{x}_m$  z aproksimacijo v obliki linearnega modela. Ker je ta model za nelinearne sisteme dober samo za področje v neposredni okolici točke  $\mathbf{x}_m$ , mu pravimo *lokalni model* (angl. *local model*, LM) in ga označimo z  $\mathcal{M}_m$ :

$$y(\mathbf{x}) = f(\mathbf{x}_m) + \boldsymbol{\theta}_m^T (\mathbf{x} - \mathbf{x}_m), \quad (4.5)$$

kjer je  $\mathbf{x}_m$  vektor regresorjev dinamičnega sistema (4.3) in so  $\boldsymbol{\theta}_m$  parametri linearnega lokalnega modela  $\mathcal{M}_m$ , identificiranega v okolici točke  $\mathbf{x}_m$ , ki ji bomo rekli tudi središče lokalnega modela:

$$\boldsymbol{\theta}_m^T = [\mathbf{a}_m^T, \mathbf{b}_m^T], \quad (4.6)$$

$$\mathbf{a}_m = \left[ \frac{\partial f}{\partial y(k-1)}, \dots, \frac{\partial f}{\partial y(k-L)} \right]_m^T, \quad (4.7)$$

$$\mathbf{b}_m = \left[ \frac{\partial f}{\partial u(k-1)}, \dots, \frac{\partial f}{\partial u(k-L)} \right]_m^T. \quad (4.8)$$

Linearne lokalne modele lahko pridobimo s katero izmed standardnih identifikacijskih metod za linearne sisteme, prim. npr. [64, 69].

Lokalni model  $\mathcal{M}_m$  sestavljata dva tipa podatkov:

1. Vrednost funkcije  $y_m = f(\mathbf{x}_m)$  (tudi izmerjena funkcijska vrednost, angl. *functional observation* [99]), ki določa točko, v okolici katere je bil identificiran linearni lokalni model  $\mathcal{M}_m$ . To je hkrati tudi delovna točka sistema, ko gre za model na ravnotežni krivulji.

2. (Parcialni) odvodi izhoda funkcije  $\frac{\partial f}{\partial x_d}$  (tudi izmerjeni odvodi<sup>2</sup>, angl. *derivative observations* [99]) v točki  $f(\mathbf{x}_m)$ , ki določajo spremembo izhoda  $f(\mathbf{x})$  glede na spremembo posameznih vhodnih regresorjev  $\mathbf{x}_m$ :

$$\boldsymbol{\theta}_m = \left[ \frac{\partial f}{\partial y^{(k-1)}}, \dots, \frac{\partial f}{\partial y^{(k-L)}}, \frac{\partial f}{\partial u^{(k-1)}}, \dots, \frac{\partial f}{\partial u^{(k-L)}} \right]_m^T.$$

Če želimo v GP model vključiti lokalne modele, moramo poleg vrednosti funkcije omogočiti še vključevanje vrednosti odvodov, ki predstavljajo parametre (koeficiente) linearnih lokalnih modelov v izbranih točkah.

#### 4.2.2 Vključevanje lokalnih modelov v GP model

Odvode v GP model lahko vključujemo, saj odvod Gaussovega procesa ostane Gaussov proces [99]. Ker pa smo uvedli nov tip podatka, se mora za ta tip podatka ustrezno spremeniti tudi kovariančna funkcija  $C(\cdot, \cdot)$ , ki vrednoti medsebojno odvisnost podatkov. Če smo predhodno potrebovali kovariančno funkcijo, ki je določala vrednost kovariance med dvema vrednostima funkcije  $y_i$  in  $y_j$ :  $\text{cov}(y_i, y_j) = C(\mathbf{x}_i, \mathbf{x}_j)$ , potrebujemo še:

1. kovariančno funkcijo, ki določa vrednost kovariance med odvodom  $\frac{\partial y_i}{\partial x_i^d}$  (parametrom linearnega modela v točki  $\mathbf{x}_i$ ) in vrednostjo funkcije  $y_j$ :

$$\text{cov}\left(\frac{\partial y_i}{\partial x_i^d}, y_j\right) = C^d(\mathbf{x}_i, \mathbf{x}_j) \quad (4.9)$$

ter

2. kovariančno funkcijo, ki določa vrednost kovariance med odvodoma  $\frac{\partial y_i}{\partial x_i^d}$  in  $\frac{\partial y_j}{\partial x_j^e}$  (parametroma linearnih modelov v točkah  $\mathbf{x}_i$  in  $\mathbf{x}_j$ ):

$$\text{cov}\left(\frac{\partial y_i}{\partial x_i^d}, \frac{\partial y_j}{\partial x_j^e}\right) = C^{de}(\mathbf{x}_i, \mathbf{x}_j). \quad (4.10)$$

V obeh primerih indeksa  $i$  in  $j$  predstavljata posamezni učni primer (vektor), indeksa  $d$  in  $e$  pa regresor, po katerem odvajamo ustrezni učni vhod.

Pri uporabi Gaussove kovariančne funkcije (2.4) za izračun kovariance med dvema vrednostima funkcije, kovarianco med odvodom in vrednostjo funkcije

<sup>2</sup>Za računanje vrednosti parcialnih odvodov izhoda po posameznih regresorjih ne potrebujemo odvodov signalov v izbranih točkah, ampak vrednosti odvodov (parametrov) dobimo z identifikacijo lokalnih modelov iz signalov.

izračunamo:

$$C_G^d(\mathbf{x}_i, \mathbf{x}_j) = \text{cov}\left[\frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, f(\mathbf{x}_j)\right] = -vw_d(x_i^d - x_j^d) \exp\left[-\frac{1}{2} \sum_{d=1}^D w_d(x_i^d - x_j^d)^2\right], \quad (4.11)$$

kovarianco med dvema odvodoma pa:

$$\begin{aligned} C_G^{de}(\mathbf{x}_i, \mathbf{x}_j) &= \text{cov}\left[\frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, \frac{\partial f(\mathbf{x}_j)}{\partial x_j^e}\right] = \\ &= vw_e(\delta_{e,d} - w_d(x_i^e - x_j^e)(x_i^d - x_j^d)) \exp\left[-\frac{1}{2} \sum_{d=1}^D w_d(x_i^d - x_j^d)^2\right], \end{aligned} \quad (4.12)$$

kjer je  $\delta_{e,d}$  Kroneckerjev operator med indeksoma  $d$  in  $e$ :

$$\delta_{e,d} = \begin{cases} 1, & e = d; \\ 0, & e \neq d. \end{cases} \quad (4.13)$$

## Sestavljanje LMGP modela

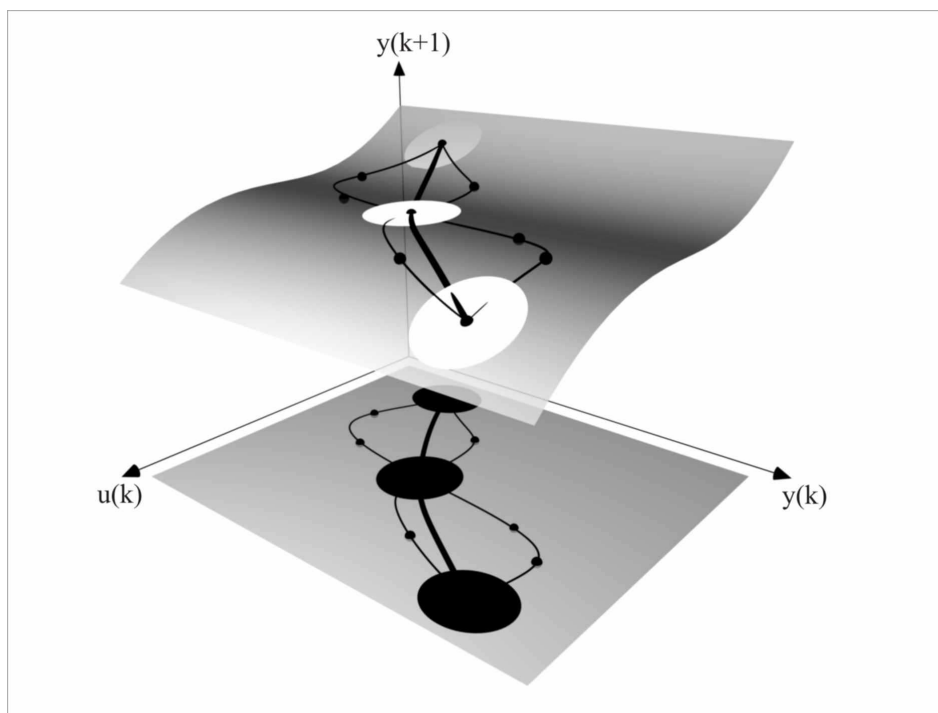
Z možnostjo vključevanja odvodov v GP model imamo na voljo vse potrebne sestavne dele za sestavo LMGP modela, ki je sestavljen iz lokalnih modelov (vrednosti funkcije in odvodov) in iz vzorcev obnašanja sistema (vrednosti funkcije). Opisana kombinacija dveh tipov podatkov je možna rešitev problema opisovanja dinamike na neravnotežnih področjih z lokalnimi modeli, ki smo ga opisali v podpoglavju 3.1 in je opisan v [35, 37, 75]:

Z lokalnimi modeli opišemo tista področja sistema, s katerih je na voljo dovolj podatkov za njihovo identifikacijo. Običajno je to v bližini ravnotežne krivulje<sup>3</sup>. Področja, kjer je podatkov za identifikacijo z LM premalo, opišemo s posameznimi vzorci vhodno/izhodnih podatkov sistema, kot smo to storili v “običajnem” GP modelu. Obe predstavitvi sistema — lokalne modele in vzorce odziva — združimo v LMGP model, kot je ilustrirano na sliki 4.3 za dinamični sistem prvega reda<sup>4</sup>. LMGP model informacijo obeh predstavitev zgladi in daje robustno napoved obnašanja sistema tudi na področju, opisanem z malo učnimi točkami.

Z vključitvijo lokalnih modelov v GP model smo vektorju učnih izhodov  $\mathbf{y}$ , ki je bil prej sestavljen samo iz vrednosti izhodov sistema, dodali vrednosti parametrov

<sup>3</sup>Če bi radi vključili lokalni model iz neravnotežnega področja, se postopek ne spremeni; tak model lahko vstavimo na enak način kot lokalni model iz ravnotežnega področja.

<sup>4</sup>Tu velja omeniti še drugo možnost modeliranja sistema [34]: področje delovanja  $\mathcal{R}^D$  razdelimo na manjša območja, ki jih nato vsakega posebej opišemo z GP modelom. Nato z uporabo variance z optimizacijo poiščemo strukturo LMN.



**Slika 4.3:** Učni izhodi LMGP modela so vrednosti parametrov lokalnih modelov *in* vzorci izhoda sistema. Na sliki je ilustriran princip metode za dinamični sistem prvega reda. Krogi predstavljajo območje veljavnosti lokalnih modelov, točke pa vzorce izhoda sistema na neravnotežnem področju.

lokalnih modelov  $\theta_m$ ,  $m = 1, \dots, n_{eq}$ , kjer je  $n_{eq}$  število lokalnih modelov. Matriki učnih vhodov  $\mathbf{X}$  so ustrezno dodane vrednosti regresorjev  $\mathbf{x}_m$  iz (4.5), ki določajo delovno točko (središča) lokalnih modelov ustreznih vrednosti parametrov.

Recimo, da poleg  $n_{eq}$  lokalnih modelov obnašanje sistema opisuje še  $n_{oeq}$  vzorcev odziva sistema. Ena izmed možnosti za sestavo vhodno/izhodnih učnih podatkov  $\mathcal{D} = \mathbf{X}, \mathbf{y}$  je:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{oeq} \\ \mathbf{X}_{eq} \\ \mathbf{X}_{eq} \\ \vdots \\ \mathbf{X}_{eq} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \mathbf{y}_{oeq} \\ \mathbf{y}_{eq} \\ \theta^1 \\ \vdots \\ \theta^D \end{bmatrix}, \quad (4.14)$$

kjer so:

$\mathbf{y}_{oeq}$  vektor želenih odzivov zunaj ravnotežne krivulje dolžine  $n_{oeq}$ ;

$\mathbf{X}_{oeq}$  vhodna matrika regresorjev velikosti  $n_{oeq} \times D$ , ki ustreza vektorju učnih izhodov  $\mathbf{y}_{oeq}$ ;

$\mathbf{y}_{eq}$  vektor vrednosti izhoda sistema dolžine  $n_{eq}$  v središčih lokalnih modelov;

$\mathbf{X}_{eq}$  matrika regresorjev velikosti  $n_{eq} \times D$ , ki ustreza vektorju učnih izhodov  $\mathbf{y}_{eq}$ ;

$\boldsymbol{\theta}^1$  vektor odvodov  $\frac{\partial f}{\partial \mathbf{y}^{(k-1)}}$  velikosti  $n_{eq} \times 1$ , ki ustreza matriki učnih vhodov  $\mathbf{X}_{eq}$   
(vektor odvodov  $\frac{\partial f}{\partial \mathbf{y}^{(k-1)}}$  vseh  $n_{eq}$  vključenih lokalnih modelov);

$\boldsymbol{\theta}^L$  vektor odvodov  $\frac{\partial f}{\partial \mathbf{y}^{(k-L)}}$  velikosti  $n_{eq} \times 1$ , ki ustreza matriki učnih vhodov  $\mathbf{X}_{eq}$ ;

$\boldsymbol{\theta}^{L+1}$  vektor odvodov  $\frac{\partial f}{\partial \mathbf{u}^{(k-1)}}$  velikosti  $n_{eq} \times 1$ , ki ustreza matriki učnih vhodov  $\mathbf{X}_{eq}$ ;

$\boldsymbol{\theta}^{2L}$  vektor odvodov  $\frac{\partial f}{\partial \mathbf{u}^{(k-L)}}$  velikosti  $n_{eq} \times 1$ , ki ustreza matriki učnih vhodov  $\mathbf{X}_{eq}$ .

Kot v enačbi (4.4), je tudi tu  $L$  red sistema in  $D = 2L$  število regresorjev. Index  $oeq$  označuje vzorčne izhode is sistema na neravnotežnem področju, indeks  $eq$  pa označuje opis obnašanja sistema s parametri lokalnih modelov na ravnotežnem področju. Velja pripomniti, da smo z vključitvijo lokalnih modelov z ravnotežne krivulje v GP model s točkami  $\mathbf{X}_{eq}, \mathbf{y}_{eq}$  avtomatsko vključili tudi statično karakteristiko sistema.

Sistem je predstavljen z  $n = n_{eq} + n_{oeq}$  funkcijskimi vrednostmi (vhodno/izhodnimi učnimi točkami) in dodatno z  $D = 2L$  vektorji odvodov  $\boldsymbol{\theta}^k$  dolžine  $n_{eq}$ , po en vektor za odvod po vsakem izmed regresorjev  $\mathbf{x}_m$ . Velikost vhodne učne matrike  $\mathbf{X}$  je tako  $(n + D n_{eq}) \times D$ , dolžina ustreznega izhodnega učnega vektorja  $\mathbf{y}$  pa  $(n + D n_{eq})$ .

Ko učne podatke  $\mathcal{D} = \mathbf{X}, \mathbf{y}$  sestavimo po opisanem principu, za kovariančno matriko  $\mathbf{K}$ , vektor kovarianc med testnim vhomom in učnimi vhodi  $\mathbf{k}(\mathbf{x})$  ter avtokovarianco testnega izhoda  $k(\mathbf{x})$  dobimo:

$$\mathbf{K} = \begin{bmatrix} [C(\mathbf{x}_i, \mathbf{x}_j)] & \left[ \text{cov}\left[f(\mathbf{x}_i), \frac{\partial f(\mathbf{x}_j)}{\partial x_j^e}\right] \right]_{e=1} & \dots & \left[ \text{cov}\left[f(\mathbf{x}_i), \frac{\partial f(\mathbf{x}_j)}{\partial x_j^e}\right] \right]_{e=D} \\ \left[ \text{cov}\left[\frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, f(\mathbf{x}_j)\right] \right]_{d=1} & \left[ \text{cov}\left[\frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, \frac{\partial f(\mathbf{x}_j)}{\partial x_j^e}\right] \right]_{d=1, e=1} & \dots & \left[ \text{cov}\left[\frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, \frac{\partial f(\mathbf{x}_j)}{\partial x_j^e}\right] \right]_{d=1, e=D} \\ \vdots & \vdots & \vdots & \vdots \\ \left[ \text{cov}\left[\frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, f(\mathbf{x}_j)\right] \right]_{d=D} & \left[ \text{cov}\left[\frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, \frac{\partial f(\mathbf{x}_j)}{\partial x_j^e}\right] \right]_{d=D, e=1} & \dots & \left[ \text{cov}\left[\frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, \frac{\partial f(\mathbf{x}_j)}{\partial x_j^e}\right] \right]_{d=D, e=D} \end{bmatrix}, \quad (4.15)$$

$$\mathbf{k}(\mathbf{x}) = \begin{bmatrix} [C(\mathbf{x}_i, \mathbf{x})] \\ \left[ \text{cov} \left[ \frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, f(\mathbf{x}) \right] \right]_{d=1} \\ \vdots \\ \left[ \text{cov} \left[ \frac{\partial f(\mathbf{x}_i)}{\partial x_i^d}, f(\mathbf{x}) \right] \right]_{d=D} \end{bmatrix}, \quad (4.16)$$

$$k(\mathbf{x}) = [C(\mathbf{x}, \mathbf{x})] = v. \quad (4.17)$$

Učenje LMGP modela je enako kot učenje običajnega GP modela. Optimirati je potrebno samo hiperparametre kovariančnih funkcij (2.4), (4.11) in (4.12). Velja opozoriti na dve stvari:

- vse tri naštete kovariančne funkcije uporabljajo *skupne* hiperparametre in
- parametri lokalnih modelov se ne optimirajo (spreminjajo), ampak obdržijo vrednosti, dobljene z identifikacijo.

GP model z vključenimi lokalnimi modeli ponuja razmeroma eleganten način združevanja informacije o obnašanju sistema v obliki lokalnih modelov in vzorcev izhodnega signala sistema. Kjer je na voljo dovolj podatkov za identifikacijo (običajno na ravnotežni krivulji), je sistem opisan z lokalnimi modeli, drugod pa z vzorci svojega odziva v odvisnosti od vrednosti regresorjev. To, da z uporabo lokalnih modelov lahko informacijo o obnašanju sistema v bližini točke  $\mathbf{x}_m$  namesto z (velikim) številom vzorcev izhoda predstavimo s parametri lokalnega modela  $\mathcal{M}_m$ , zmanjša učno množico in potencialno zmanjša tudi računsko zahtevnost metode predvsem v fazi optimizacije hiperparametrov.

Izmerjeni parametri lokalnih modelov lahko vsebujejo tudi informacijo o šumu v obliki variance parametrov, ki jo lahko dodamo v šumni del kovariančne matrike [99]. Če te informacije nimamo, parametre šumnega procesa optimiramo kot za "običajni" GP model.

Moramo paziti, da imajo vsi lokalni modeli enaka notranja stanja. Stanja sistema so namreč izražena z vrednostmi regresorjev, ki nastopajo kot vhod v LMGP model in prek katerih se izražajo vrednosti parametrov lokalnih modelov. V našem primeru smo uporabili kar vrednosti regresorjev (4.4), lahko pa si izberemo drugačno predstavitev.

Dinamika LMGP modela v neravnotežnih področjih je opisana z vzorci izhodnega signala sistema in tako predstavlja globalno obnašanje. Po drugi strani vključeni

lokalni modeli v ravnotežnih območjih opisujejo lokalno dinamiko sistema, in ker se parametri identificiranih lokalnih modelov z optimizacijo hiperparametrov LMGP modela ne spremenijo, lahko trdimo, da dinamika sistema na tem področju ostane ustrezno opisana. Prav tako je z uporabo LMGP modela manj težavno razdeliti področje delovanja med lokalne modele kot pri LMN, saj lokalne modele postavimo le na ravnotežno krivuljo.

Poleg odprave nekaterih težav LMN daje napovedovanje z LMGP modelom poleg napovedi srednje vrednosti izhoda tudi *mero zaupanja* v to napoved, katere uporabnost smo spoznali v prejšnjih poglavjih.

### 4.2.3 Napovedovanje z uporabo LMGP modela in njegova simulacija

Napovedovanje LMGP modela pri novem vhodu  $\mathbf{x}$  ostane enako kot pri “običajnem” GP modelu. Srednja vrednost  $\mu(\mathbf{x})$  in varianca  $\sigma^2(\mathbf{x})$  napovedane izhodne Gaussove porazdelitve sta:

$$\mu(\mathbf{x}) = \mathbf{k}(\mathbf{x})^T \mathbf{K}^{-1} \mathbf{y}, \quad (4.18)$$

$$\sigma^2(\mathbf{x}) = k(\mathbf{x}) - \mathbf{k}(\mathbf{x})^T \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}). \quad (4.19)$$

### Napovedovanje z uporabo LMGP modela pri porazdelitvi na vhodu v GP model

Ideja napovedovanja pri porazdeljenem vhodu je za LMGP model enaka kot za “običajni” GP model (podpoglavje 3.5.2). Namesto samo srednje vrednosti se iz izhoda na vhod prenese celotna porazdelitev (ta je predpostavljeno normalna, pri čemer gre za aproksimacijo). Tako izveden LMGP model, enako kot “običajni” GP model, bolj natančno izraža stopnjo zaupanja v napovedano vrednost.

### Simulacija LMGP modela

Postopek simulacije LMGP modela je tako za simulacijo brez propagacije negotovosti kot za simulacijo s propagacijo negotovosti enak kot za “običajni” GP model in je opisan v razdelku 3.5.

Za propagacijo negotovosti imamo na voljo analitično in numerično aproksimacijo napovedovanja LMGP modela. Numerična propagacija negotovosti se razen v modelu v ničemer ne razlikuje od numerične propagacije pri “običajnem” GP modelu.

Analitična propagacija negotovosti temelji na ločevanju obeh tipov podatkov, ki sestavljata LMGP model. Postopek je opisan v [27, 46], tu ga zaradi dolžine ne bomo navajali. Razlika se pojavi tudi pri računanju vrednosti križnih kovarianc med izhodi, ki se spremeni [27].

#### 4.2.4 Vrednotenje

Vrednotenje celotnega LMGP modela se v ničemer ne razlikuje od vrednotenja “običajnega” GP modela (podpoglavje 3.6), tudi tu lahko uporabimo napovedane variance GP modela.

Kar LMGP model loči od “običajnega”, so vključeni lokalni modeli. Te moramo ovrednotiti vsakega posebej za svoje področje še pred vključevanjem.

#### 4.2.5 Primer identifikacije dinamičnega sistema drugega reda z LMGP modelom

Identifikacijo z LMGP modelom bomo ilustrirali na dinamičnem sistemu drugega reda (4.2) na področju, omejenem z vhodnim signalom  $u$  (med  $u_{\min} = -2$  in  $u_{\max} = 4$ ) in varianco belega Gaussovega šuma  $\epsilon(k)$  na izhodu  $\sigma^2 = 4 \cdot 10^{-4}$ . Delovanje identificiranega modela bomo ovrednotili s simulacijo.

Isti primer, a z drugačnimi učnimi podatki in podatki za vrednotenje, smo uporabili že v [4]. S spremembo podatkov smo želeli doseči večjo ilustrativnost primera.

### Identifikacija

Kot je bilo povedano, LMGP model sestavljajo:

- lokalni modeli, ki opisujejo obnašanje sistema v okolici svojih središč na ravnotežni krivulji, in sicer v obliki parametrov kot učnih podatkov ter
- vzorci izhoda sistema, ki opisujejo obnašanje sistema na področjih, ki niso opisana z lokalnimi modeli; navadno so to prehodna območja med ravnotežnimi stanji.

Za pridobitev teh dveh tipov informacije o sistemu potrebujemo dva tipa meritev. Da pridobimo podatke, ki opisujejo neravnotežna področja sistema, moramo sistem vzbujati tako, da vzorci regresorjev, ki so uporabljeni kot vhodni del

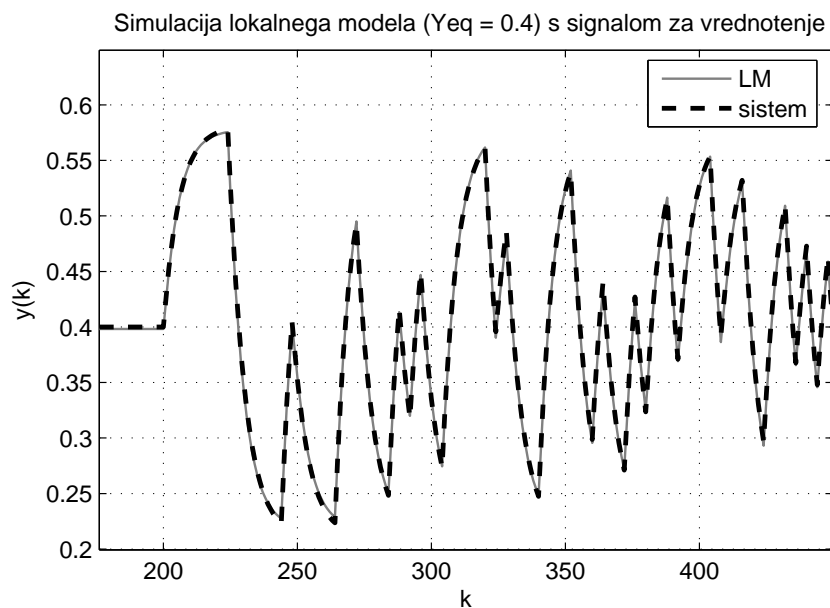
učnih točk, pokrijejo čim večje področje delovanja sistema  $\mathcal{R}^D$ . V našem primeru smo za vzbujevalni signal uporabili diskretni psevdo-naključni signal, ki smo ga dobili iz psevdo-naključnega binarnega signala (angl. *pseudo-random binary signal*, PRBS) [69], katerega amplituda signala  $u$  je ob spremembi lahko zavzela katerokoli vrednost med  $u_{\min}$  in  $u_{\max}$ . Učne točke dobimo z vzorčenjem vhoda  $u$  in izhoda  $y$  sistema.

Obnašanje modela v bližini ravnotežne krivulje je opisano z lokalnimi modeli. Njihove parametre dobimo tako, da sistem najprej vzbujamo s konstantnim vhodnim signalom in počakamo, da se ustali v ravnotežni točki. Nato vходу  $u$  dodamo PRBS z majhno amplitudo  $\Delta U$ , da dobimo dinamični odziv sistema v okolici delovne točke. Amplitudo  $\Delta U$  izberemo tako, da je z lokalnimi modeli pokrit čim večji del ravnotežne krivulje. Pri tem pazimo, da  $\Delta U$  ni premajhen zaradi razmerja signal/šum meritev na izhodu in hkrati ne prevelik, da identificirani parametri še vedno opisujejo lokalno dinamiko sistema. Lokalne modele identificiramo z uporabo katere izmed standardnih identifikacijskih metod [64, 69].

Za opis dinamike na ravnotežnem področju sistema (4.2) smo identificirali pet približno enakomerno razporejenih lokalnih modelov. Njihova središča vidimo na sliki 4.2 zgoraj. Izbran je bil PRBS s taktom  $T_{sw} = 4$  korakov in velikostjo vzbujanja  $\Delta U = 0.3$ , da so bili lokalni modeli dobro identificirani kljub šumu. Za identifikacijo je bila uporabljena *metoda pomožnih spremenljivk* (angl. *Instrumental Variables*, IV) [64, 69].

Primer odziva identificiranega lokalnega modela v ravnotežni točki  $(U_{eq}, Y_{eq}) = (0.415, 0.4)$  na signalu, ki ni bil uporabljen za identifikacijo (brez šuma), vidimo na sliki 4.4 in lahko opazimo, da model popolno opiše dinamiko sistema. Seveda se moramo zavedati, da imamo opravka z idealnim sistemom znanega reda.

Vsak izmed petih lokalnih modelov k učnim podatkom prispeva eno vrednost funkcije (središče lokalnega modela) in štiri vrednosti odvodov, ki jih predstavljajo identificirani parametri lokalnih modelov, po en za vsak regresor. Skupno s štirinajstimi vzorčenimi točkami odziva, ki opisujejo obnašanje sistema v neravnotežnem področju, učno množico LMGP modela sestavlja 39 točk. Ocenjena varianca parametrov lokalnih modelov, ki smo jo dobili v postopku identifikacije lokalnih modelov, je bila dodana ustreznim elementom kovariančne matrike (4.15).



**Slika 4.4:** Primer odziva identificiranega lokalnega modela s središčem v ravnotežni točki  $(U_{eq}, Y_{eq}) = (0.415, 0.4)$

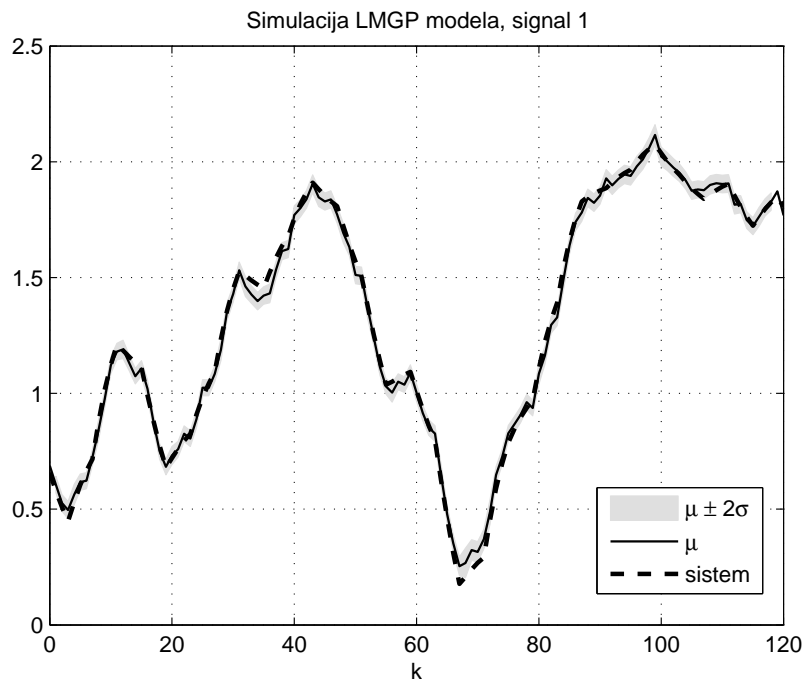
### Vrednotenje

Sestavljen in naučen LMGP model smo ovrednotili s simulacijo na dveh vhodnih signalih, ki nista bila uporabljena za identifikacijo, a sta imela enake lastnosti, tj. takt  $T_{sw} = 4$  korake in vhod  $u$  med  $u_{\min} = -2$  in  $u_{\max} = 4$ :

1. prvi signal za vrednotenje, pri katerem je bila največja sprememba vhodnega signala omejena, da bi se LMGP pomikal samo skozi bolj opisana področja in
2. drugi signal za vrednotenje, pri katerem sprememba vhoda ni bila omejena in se je LMGP model premikal tudi skozi področje, kjer ni bilo učnih podatkov.

Dva različna signala za vrednotenje smo izbrali, da lahko boljše ilustriramo lastnosti GP modela. Poglejmo si najprej rezultate vrednotenja s prvim signalom. Rezultati simulacije s tem vhodnim signalom so prikazani na sliki 4.5, ustrezna napaka simulacije skupaj s 95% pasom zaupanja v napoved LMGP modela pa je prikazana na sliki 4.6. Naj opozorimo, da je zaradi preglednosti na teh dveh slikah predstavljen samo del celotnih rezultatov. Avtokorelacija napake simulacije  $\Phi_{ee}$  in križna korelacija med vhomom in napako simulacije  $\Phi_{ue}$  sta prikazani na sliki 4.7. Iz slik 4.5–4.7 lahko sklepamo, da je opis obnašanja sistema v neravnotežnih

področjih zadovoljiv, čeprav je bilo za njihov opis na voljo le štirinajst učnih točk. Odziv bi seveda lahko izboljšali z dodajanjem vzorcev med učne točke.

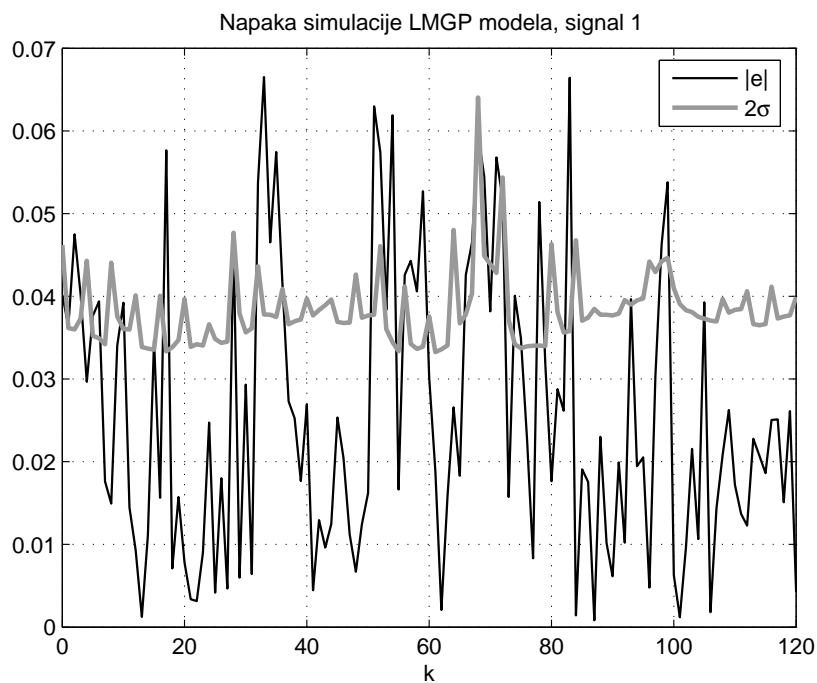


**Slika 4.5:** Simulacija identificiranega LMGP modela s prvim signalom za vrednotenje

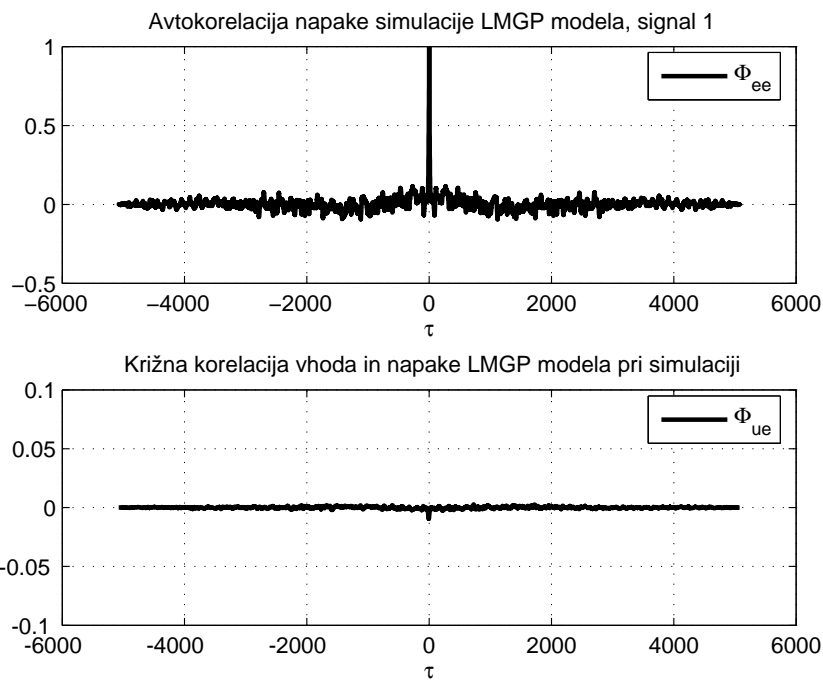
Del rezultatov simulacije na drugem signalu za vrednotenje je prikazan na slikah 4.8 in 4.9; na prvi sliki je napovedan izhod LMGP modela skupaj s 95% pasom zaupanja v napoved primerjan z izhodom sistema, na drugi pa je s 95% pasom zaupanja primerjana napaka simulacije. Ti rezultati kažejo, da se varianca napovedi poveča na področjih delovanja sistema, ki niso zadovoljivo opisana, kar je lepo vidno tudi iz primerjave napak in napovedanih standardnih odstopanj modela med simulacijo na slikah 4.6 in 4.9.

Za vrednotenje rezultatov simulacije z obema vhodnima signaloma so bile uporabljene tudi cenilke povprečna kvadratna napaka SE (3.20), povprečna relativna kvadratna napaka MRSE (3.21) in logaritem napovedane gostote napake LD (3.22), rezultati vrednotenja so zbrani v tabeli 4.1. Rezultati so zadovoljivi, opazno je boljše obnašanje modela, ko ga simuliramo na prvem signalu za vrednotenje, ker se model premika skozi boljše opisana področja.

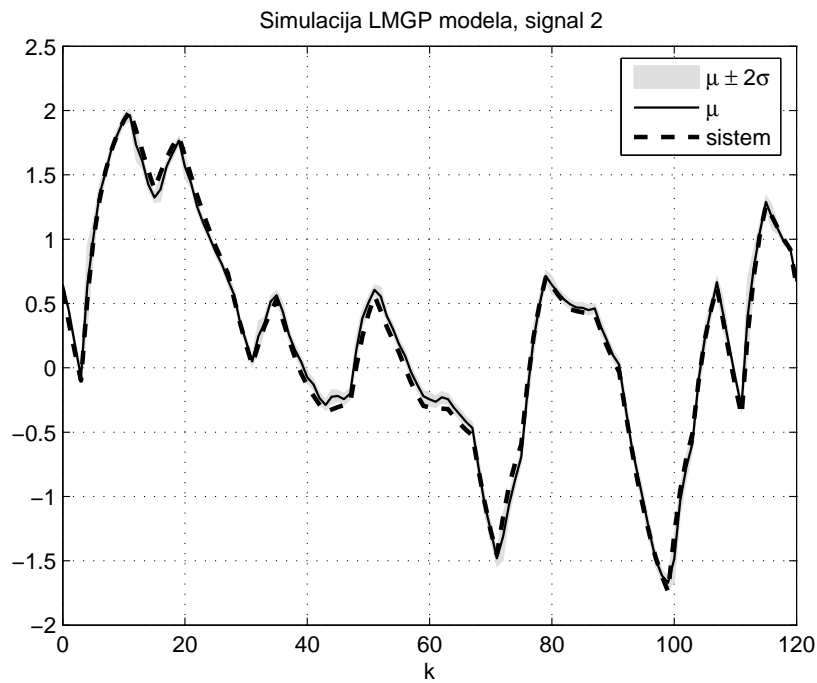
\* \* \*



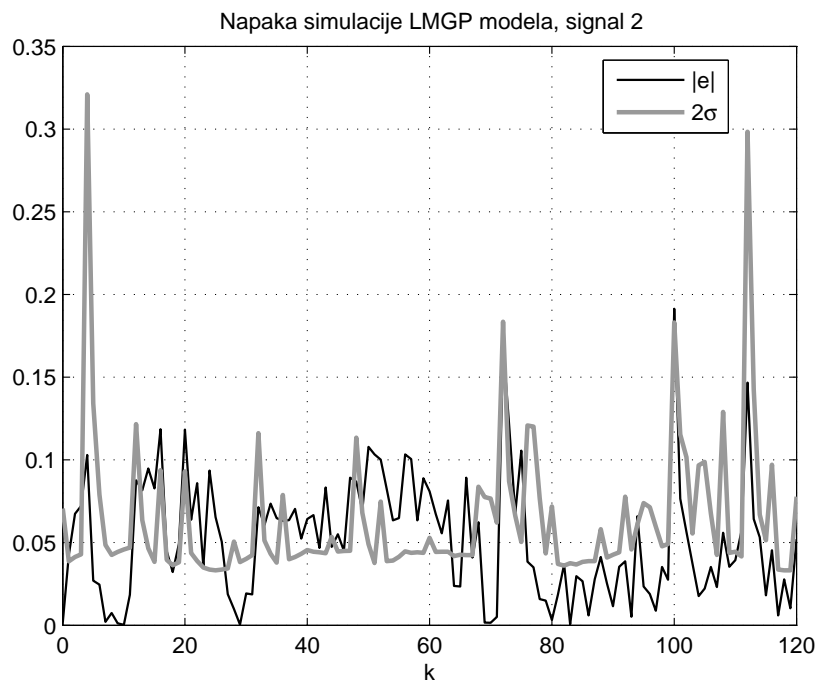
**Slika 4.6:** Absolutna napaka LMGP modela pri simulaciji s prvim signalom za vrednotenje skupaj s 95% pasom zaupanja



**Slika 4.7:** Avtokorelacija napake simulacije  $\Phi_{ee}$  (zgoraj) in križna korelacija med vhodom v LMGP model in napako simulacije  $\Phi_{ue}$  (spodaj) za prvi signal za vrednotenje



**Slika 4.8:** Simulacija identificiranega LMGP modela z drugim signalom za vrednotenje



**Slika 4.9:** Absolutna napaka LMGP modela pri simulaciji z drugim signalom za vrednotenje skupaj s 95% pasom zaupanja

**Tabela 4.1:** Vrednost cenilk SE, MRSE in LD pri simulaciji LMGP modela na obeh signalih za vrednotenje

Cenilka	Prvi signal	Drugi signal
SE	$1.7 \cdot 10^{-3}$	$3.2 \cdot 10^{-3}$
MRSE	0.037	0.052
LD	-1.05	-0.47

Z uporabo LMGP modela smo zmanjšali probleme, ki bi jih imeli pri uporabi LMN:

- Ni potrebna identifikacija lokalnih modelov po vsem območju delovanja sistema, lahko jih identificiramo samo v ravnotežnem področju. Zato je zmanjšan tudi problem izbire lege in števila lokalnih modelov.
- Obnašanje sistema v neravnotežnem področju je predstavljeno samo s posameznimi vzorci izhodnega signala, na vmesnih področjih LMGP model pri napovedovanju gladko interpolira.
- Predstavitev dinamičnega obnašanja modela ostane taka, kot je bila pred učenjem, saj se lokalni modeli z optimizacijo hiperparametrov LMGP modela ne spremenijo. Tako izgine potreba po odločanju, ali optimirati globalno ali lokalno obnašanje sistema.
- Ni problema izbire razvrstilnih spremenljivk. Vrednosti hiperparametrov kovariančne funkcije lahko uporabimo kot mero pomembnosti ustreznih regresorjev.

V primerjavi z “običajnim” GP modelom LMGP model

- uporablja predznanje o sistemu v obliki (že identificiranih) lokalnih modelov.
- Uporaba lokalnih modelov zmanjša število učnih točk in posledično lahko zmanjša časovno zahtevnost učenja globalnega modela in njegovega napovedovanja.
- Uporaba lokalnih modelov prav tako poveča transparentnost modela, saj je dinamično obnašanje sistema na ravnotežni krivulji opisano z lokalnimi modeli, vendar
- se pojavi potreba po identifikaciji lokalnih modelov, kar po drugi strani podaljša postopek identifikacije.

Uporaba LMGP modela (konstrukcija vhodno/izhodnih učnih podatkov, optimizacija hiperparametrov, vrednotenje itd.) ni nič težja od uporabe “običajnega” GP modela, so pa izrazi za kovariančno funkcijo bolj zapleteni, kar nekoliko vpliva na računsko zahtevnost.

Iz naštetega sledi, da je LMGP model koristno uporabiti tam, kjer so lokalni modeli za sistem že identificirani oz. je to lahko storiti, hkrati pa je na voljo malo podatkov o obnašanju sistema na neravnotežnih področjih.

### 4.3 GP model z vnaprej določeno strukturo

Do sedaj smo že dokaj podrobno predstavili GP model. Kljub dobrim lastnostim modela, koristnim tudi pri opisovanju nelinearnih dinamičnih sistemov, nekatere njegove lastnosti omejujejo možnosti za njegovo uporabo. Taka lastnost je *neparametričnost* modela, ki na primer omejuje izbiro postopkov za načrtovanje vodenja, kadar želimo model uporabiti v ta namen.

Z uporabo GP modela za načrtovanje vodenja je povezana tudi naslednja opisana lastnost. Ker se z dodajanjem novih podatkov model povečuje, rekurzivna identifikacija ni praktična, dokler ne bo odkrita možnost posodobitve modela brez povečanja njegovega obsega. Poleg tega je potrebno po vsaki posodobitvi modela tega še enkrat učiti, kar je zelo zamuden postopek. Opisano onemogoča uporabo GP modelov za adaptivno vodenje. Edina praktična možnost za načrtovanje vodenja je načrtovanje prediktivnega vodenja na podlagi modela (angl. *model predictive control*, MPC) [62]. Za uporabo katere izmed drugih oblik vodenja, pri kateri na primer parametre regulatorja dobimo z zlivanjem parametrov lokalnih regulatorjev, potrebujemo parametrični model sistema.

Za opis nelinearnega dinamičnega sistema s parametričnim modelom obstaja več možnosti. Če poznamo strukturo opisanega dinamičnega sistema, je potrebno iz podatkov optimirati samo parametre modela iste strukture. Če tega znanja nimamo, je naslednja možnost, da sistem opišemo z mrežo lokalnih modelov (podpoglavje 3.1). Tak način je še posebej zanimiv, če gradimo model za uporabo v načrtovanju vodenja, saj lahko parametre regulatorja načrtujemo za vsak posamezni lokalni model posebej [32]. Ker je nepraktično opisati vsako točko s področja delovanja sistema s svojim lokalnim modelom, moramo za določanje parametrov lokalnega modela v neopisanih točkah zlivati parametre okoliških lokalnih modelov. To pripelje do problemov LMN, opisanih v podpoglavju 3.1. Kot smo že omenili, lahko nekatere od teh problemov zmanjšamo z uporabo hitrostne linearizacije (VBL) [53]. Ta omogoča, da neznani dinamični sistem z lokalnimi modeli predstavimo ne le na ravnotežnih področjih, ampak na celotnem delovnem območju, pri čemer pa se še vedno pojavlja problem, kako izvesti zlivanje lokalnih modelov in kako izbrati razvrstilne spremenljivke. Tukaj vidimo priložnost za uporabo GP modelov.

V tem podpoglavju bo predstavljen model na osnovi Gaussovih procesov z vnaprej določeno strukturo (angl. *Fixed-Structure GP model*, FSGP model), ki predstavlja *parametrični* pristop k identifikaciji nelinearnih dinamičnih sistemov. Predlagamo model z vnaprej določeno, v našem primeru linearno, strukturo in spremenljivimi parametri (angl. *linear parameter varying model*, LPV model) [14, 94], pri čemer za določanje trenutnih vrednosti spremenljivih parametrov modela, odvisnih od področja delovanja, uporabljamo GP modele. Kot bomo videli,

je prednost tega pristopa dodatna informacija o meri zaupanja v napovedane parametre modela, poleg tega ima model zmožnost olajšati izbiro razvrstilnih spremenljivk.

Osnovna ideja je, da predpostavimo strukturo modela, npr. linearno, potem pa sistem na različnih delovnih področjih opišemo s podmodeli predpostavljene strukture. Iz vrednosti parametrov, v našem primeru lineariziranih, lokalnih podmodelov poiščemo GP modele, ki opisujejo posamezne spremenljive parametre izbrane strukture. Tako dobimo model sistema s strukturo linearnega modela in GP modelom za vsak parameter v tej strukturi posebej.

### 4.3.1 Modeliranje

Pristop bomo temeljili na parametrih identificiranih diskretnih lokalnih modelov, zato predstavimo model nelinearnega dinamičnega sistema kot diskretni sistem:

$$\begin{aligned}\mathbf{x}(k+1) &= F(\mathbf{x}(k), u(k)) \\ y(k) &= G(\mathbf{x}(k), u(k)),\end{aligned}\tag{4.20}$$

kjer je  $\mathbf{x}(k) = [x(k) \dots x(k-L+1)]^T$  vektor stanj sistema in  $u(k)$  vhod v sistem  $u$  v koraku  $k$ . Čas vzorčenja  $T$  izberemo dovolj majhen, da opis (4.20) ujame vso zanimivo nelinearno dinamiko originalnega sistema, red  $L$  modela (4.20), ki določa število stanj v vektorju  $\mathbf{x}(k)$ , pa je bil izbran v skladu s Takensovim teoremom (Dodatek D). Ta sistem želimo opisati z linearnim modelom s spremenljivimi parametri. Da bi lahko uporabili nekatera dognanja, ki veljajo za zvezne sisteme, smo diskretni model sistema ponazorili v zveznem prostoru.

*Diskretni sistem* (4.20) lahko z uporabo zadrževalnika ničtega reda (angl. *zero-order hold*, ZOH) kot v [93] ponazorimo v *zveznem prostoru* v obliki zapisa zveznih signalov z zakasnitvami:

$$\begin{aligned}\mathbf{x}(t+T) &= F(\mathbf{x}(t), u(t)) \\ y(t) &= G(\mathbf{x}(t), u(t)),\end{aligned}\tag{4.21}$$

pri čemer so vrednosti stanj  $\mathbf{x}$ , izhoda  $y$  in vhoda  $u$  diskretnega in zveznega sistema v trenutkih vzorčenja,  $t = kT$ ,  $k = 1, 2, \dots$ , enake. To storimo zaradi tega, ker bomo v nadaljevanju uporabili nekatera orodja, ki veljajo le za zvezne sisteme. Ker je opisan sistem zvezen in je vhod  $u(t)$  med časi  $t = kT$  in  $t = (k+1)T$  zaradi uporabe ZOH člana konstanten, lahko predpostavimo, da se sistem med točkami vzorčenja premika gladko.

Zvezni sistem (4.21) lahko izrazimo v obliki razširjene lokalne linearne ekvivalence (angl. *extended local linear equivalence*, ELLE) [55]:

$$\begin{aligned}\mathbf{x}(t+T) &= \tilde{\mathbf{A}}\mathbf{x}(t) + \tilde{\mathbf{B}}u(t) + \mathbf{f}(\boldsymbol{\rho}), \\ y(t) &= \tilde{\mathbf{C}}\mathbf{x}(t) + \tilde{\mathbf{D}}u(t) + \mathbf{g}(\boldsymbol{\rho}),\end{aligned}\quad (4.22)$$

kjer sta  $\mathbf{x}(t) \in \mathcal{R}^L$ ,  $u(t) \in \mathcal{R}$  in so  $\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}, \tilde{\mathbf{D}}$  ustrezno velike konstantne matrike<sup>5</sup>. Nelinearni funkciji  $\mathbf{f}(\cdot)$  in  $\mathbf{g}(\cdot)$  sta odvisni od vektorja razvrstilnih spremenljivk  $\boldsymbol{\rho}(t) = \boldsymbol{\rho}(\mathbf{x}(t), u(t)) \in \mathcal{R}^q$ ,  $q \leq L + 1$ , ki izraža nelinearno odvisnost dinamike sistema od njegovega stanja in vhoda s konstantnima  $\nabla_{\mathbf{x}}\boldsymbol{\rho}$ ,  $\nabla_u\boldsymbol{\rho}$  [55].

Prvo enačbo iz (4.22) lahko, relativno glede na delovno točko  $\boldsymbol{\rho}(t) = \boldsymbol{\rho}_0$ , kjer velja  $\mathbf{x} = \mathbf{x}_0$  in  $u = u_0$ , zapišemo z uporabo:

$$\delta\mathbf{x}(t) = \mathbf{x}(t) - \mathbf{x}_0 \quad (4.23)$$

$$\delta u(t) = u(t) - u_0 \quad (4.24)$$

kot:

$$\mathbf{x}(t+T) = \tilde{\mathbf{A}}(\mathbf{x}_0 + \delta\mathbf{x}(t)) + \tilde{\mathbf{B}}(u_0 + \delta u(t)) + \mathbf{f}(\boldsymbol{\rho}_0). \quad (4.25)$$

Pri predpostavki o lokalni linearnosti v bližini delovne točke  $\boldsymbol{\rho}_0$  sledi:

$$\mathbf{x}(t+T) - \mathbf{x}_0 + \mathbf{x}_0 = \tilde{\mathbf{A}}\mathbf{x}_0 + \tilde{\mathbf{A}}\delta\mathbf{x}(t) + \tilde{\mathbf{B}}u_0 + \tilde{\mathbf{B}}\delta u(t) + \mathbf{f}_0 + \mathbf{f}_{\mathbf{x}_0}\delta\mathbf{x}(t) + \mathbf{f}_{u_0}\delta u(t) \quad (4.26)$$

$$\delta\mathbf{x}(t+T) + \mathbf{x}_0 = \left(\tilde{\mathbf{A}}\mathbf{x}_0 + \tilde{\mathbf{B}}u_0 + \mathbf{f}_0\right) + \left(\tilde{\mathbf{A}} + \mathbf{f}_{\mathbf{x}_0}\right)\delta\mathbf{x}(t) + \left(\tilde{\mathbf{B}} + \mathbf{f}_{u_0}\right)\delta u(t) \quad (4.27)$$

$$\delta\mathbf{x}(t+T) = (\mathbf{F}_0 - \mathbf{x}_0) + \left(\tilde{\mathbf{A}} + \mathbf{f}_{\mathbf{x}_0}\right)\delta\mathbf{x}(t) + \left(\tilde{\mathbf{B}} + \mathbf{f}_{u_0}\right)\delta u(t), \quad (4.28)$$

pri čemer so  $\mathbf{F}_0 = F(\mathbf{x}_0, u_0)$ ,  $\mathbf{f}_0 = f(\mathbf{x}_0, u_0)$ ,  $\mathbf{f}_{\mathbf{x}_0} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_0, u_0)$  in  $\mathbf{f}_{u_0} = \frac{\partial \mathbf{f}}{\partial u}(\mathbf{x}_0, u_0)$ .

Z odvajanjem enačbe (4.28) po času<sup>6</sup> dobimo:

$$\dot{\mathbf{x}}(t+T) = \left(\tilde{\mathbf{A}} + \mathbf{f}_{\mathbf{x}_0}\right)\dot{\mathbf{x}}(t) + \left(\tilde{\mathbf{B}} + \mathbf{f}_{u_0}\right)\dot{u}(t) \quad (4.29)$$

$$\dot{\mathbf{x}}(t+T) = \mathbf{A}(\boldsymbol{\rho}_0)\dot{\mathbf{x}}(t) + \mathbf{B}(\boldsymbol{\rho}_0)\dot{u}(t), \quad (4.30)$$

pri čemer sta  $\mathbf{A}(\boldsymbol{\rho}_0) = (\tilde{\mathbf{A}} + \mathbf{f}_{\mathbf{x}_0})$  in  $\mathbf{B}(\boldsymbol{\rho}_0) = (\tilde{\mathbf{B}} + \mathbf{f}_{u_0})$ . Pri istih pogojih lahko na enak način drugo enačbo iz (4.22) zapišemo kot:

$$\dot{y}(t) = \mathbf{C}(\boldsymbol{\rho}_0)\dot{\mathbf{x}}(t) + \mathbf{D}(\boldsymbol{\rho}_0)\dot{u}(t), \quad (4.31)$$

<sup>5</sup>Tu bomo obravnavali samo sistem z enim vhomom  $u$  in enim izhodom  $y$ , zato sta matriki  $\tilde{\mathbf{B}}$  in  $\tilde{\mathbf{C}}$  pravzaprav vektorja, matrika  $\tilde{\mathbf{D}}$  pa skalar. Kljub temu jih bomo pisali kot matrike.

<sup>6</sup>Naj opozorimo, da je  $\delta\mathbf{x}(t)$  samo za  $\mathbf{x}_0$  premaknjen vektor  $\mathbf{x}(t)$  in velja:  $\frac{d(\delta\mathbf{x}(t))}{dt} = \frac{d(\mathbf{x}(t))}{dt}$ . Tako oznako smo uporabili, da nakažemo veljavnost izpeljave za  $\mathbf{x}(t)$  v okolici delovne točke, določene z  $\boldsymbol{\rho}_0$ . Enaka je povezava med  $\delta u(t)$  in  $u(t)$ .

pri čemer sta  $\mathbf{C}(\boldsymbol{\rho}_0) = (\tilde{\mathbf{C}} + \mathbf{g}_{\mathbf{x}0})$  in  $\mathbf{D}(\boldsymbol{\rho}_0) = (\tilde{\mathbf{D}} + \mathbf{g}_{u0})$ , z  $\mathbf{g}_{\mathbf{x}0} = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}_0, u_0)$  in  $\mathbf{g}_{u0} = \frac{\partial \mathbf{g}}{\partial u}(\mathbf{x}_0, u_0)$ .

Na opisani način lahko predstavimo delovanje sistema v kateri koli točki, določeni z vektorjem razvrstilnih spremenljivk  $\boldsymbol{\rho}$ , z vektorjem stanj  $\mathbf{x}(t)$  in vhomom  $u(t)$ . Sistem (4.22) lahko zapišemo:

$$\dot{\mathbf{x}}(t+T) = \mathbf{A}(\boldsymbol{\rho})\dot{\mathbf{x}}(t) + \mathbf{B}(\boldsymbol{\rho})\dot{u}(t) \quad (4.32)$$

$$\dot{y}(t) = \mathbf{C}(\boldsymbol{\rho})\dot{\mathbf{x}}(t) + \mathbf{D}(\boldsymbol{\rho})\dot{u}(t), \quad (4.33)$$

pri čemer so matrike

$$\mathbf{A}(\boldsymbol{\rho}) = \tilde{\mathbf{A}} + \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}(t), u(t)), \quad (4.34)$$

$$\mathbf{B}(\boldsymbol{\rho}) = \tilde{\mathbf{B}} + \frac{\partial \mathbf{f}}{\partial u}(\mathbf{x}(t), u(t)), \quad (4.35)$$

$$\mathbf{C}(\boldsymbol{\rho}) = \tilde{\mathbf{C}} + \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}(t), u(t)) \quad (4.36)$$

$$\mathbf{D}(\boldsymbol{\rho}) = \tilde{\mathbf{D}} + \frac{\partial \mathbf{g}}{\partial u}(\mathbf{x}(t), u(t)) \quad (4.37)$$

odvisne od vektorja razvrstilnih spremenljivk  $\boldsymbol{\rho}(t)$ .

V vsaki točki  $\boldsymbol{\rho}(t)$  lahko sistem, opisan z enačbama (4.32) in (4.33), predstavimo tudi v obliki linearnega lokalnega modela s parametri  $\boldsymbol{\theta}(\boldsymbol{\rho}(t))$ . Ker sistema ne moremo identificirati v vseh potencialnih točkah, določenih z  $\boldsymbol{\rho}$ , je na območjih med točkami potrebno parametre lokalnih modelov zlivati. Pri tem je treba paziti, da je transformacija opisa iz (4.32)-(4.33) v lokalni model za vse delovne točke enaka, kot je to za zvezne sisteme opisano v [55]. Povedano drugače: isti parameter vseh lokalnih modelov mora izražati vpliv istega regresorja.

V disertaciji so za koordinate uporabljene pretekle vrednosti vhoda in izhoda, na voljo pa so tudi druge možnosti. V našem primeru je posamezni lokalni model  $\mathcal{M}_m$  določen kot v podpoglavju 4.2, enačba (4.5), vektor parametrov posameznega lokalnega modela  $\boldsymbol{\theta}_m$  pa je sestavljen kot:

$$\boldsymbol{\theta}_m = [-a_1^m \cdots -a_L^m \ b_1^m \cdots b_L^m], \quad (4.38)$$

kjer je  $L$  red modela. Lokalni model  $\mathcal{M}_m$  lahko z matrikami zapišemo kot [30]:

$$\mathbf{A}^m = \begin{bmatrix} -a_1^m & 1 & 0 & \cdots & 0 \\ -a_2^m & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ & & & & 1 \\ -a_L^m & 0 & \cdots & & 0 \end{bmatrix} \quad \mathbf{B}^m = \begin{bmatrix} b_1^m \\ b_2^m \\ \vdots \\ b_L^m \end{bmatrix} \quad (4.39)$$

$$\mathbf{C}^m = [1 \ 0 \ \cdots \ 0] \quad \mathbf{D}^m = [0].$$

Vrednosti posameznih parametrov lokalnih modelov, tj. istovrstni koeficienti vhodno/izhodnega modela, določajo vpliv istega regresorja na izhod sistema. Vsak parameter lahko predstavimo z neko funkcijo, odvisno od vrednosti vektorja razvrstilnih spremenljivk  $\boldsymbol{\rho} = \boldsymbol{\rho}(t)$  in jo opišemo z GP modelom. Spremenljivi parametri enačb (4.32) in (4.33) so torej:

$$\mathbf{A}(\boldsymbol{\rho}) = \begin{bmatrix} -a_1(\boldsymbol{\rho}) & 1 & 0 & \cdots & 0 \\ -a_2(\boldsymbol{\rho}) & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ -a_L(\boldsymbol{\rho}) & 0 & \cdots & & 1 \\ & & & & 0 \end{bmatrix} \quad \mathbf{B}(\boldsymbol{\rho}) = \begin{bmatrix} b_1(\boldsymbol{\rho}) \\ b_2(\boldsymbol{\rho}) \\ \vdots \\ b_L(\boldsymbol{\rho}) \end{bmatrix} \quad (4.40)$$

$$\mathbf{C} = [1 \ 0 \ \cdots \ 0] \quad \mathbf{D} = [0],$$

kjer so parametri  $\boldsymbol{\theta}(\boldsymbol{\rho}) = [-a_1(\boldsymbol{\rho}) \cdots -a_L(\boldsymbol{\rho}) \ b_1(\boldsymbol{\rho}) \cdots b_L(\boldsymbol{\rho})]$  ponazorjeni z GP modeli. Za opis vseh parametrov potrebujemo toliko GP modelov, kot je parametrov lokalnih modelov v  $\boldsymbol{\theta}(\boldsymbol{\rho})$ , ti GP modeli pa predstavljajo spremenljive parametre FSGP modela.

Kot vhod v GP modele nastopa vektor razvrstilnih spremenljivk  $\boldsymbol{\rho} = \boldsymbol{\rho}(t)$ , kot izhodi GP modelov pa nastopajo vrednosti ustreznih nekonstantnih parametrov  $\boldsymbol{\theta} = \boldsymbol{\theta}(\boldsymbol{\rho})$ . Učni podatki za GP modele so parametri identificiranih lokalnih modelov. Pri vhodnem vektorju  $\boldsymbol{\rho}$  so izhodi GP modelov napovedane vrednosti parametrov  $\hat{\boldsymbol{\theta}} = \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\theta}}(\boldsymbol{\rho}), \boldsymbol{\sigma}_{\boldsymbol{\theta}}^2(\boldsymbol{\rho}))$ , izražene s srednjimi vrednostmi  $\mu_{\theta_i}(\boldsymbol{\rho})$  in ustreznimi variancami  $\sigma_{\theta_i}^2(\boldsymbol{\rho})$ ,  $i = 1, \dots, 2L$ , kjer je z indeksom  $i$  označena napoved  $i$ -tega parametra iz vektorja parametrov lokalnega modela  $\boldsymbol{\theta}$ . Vektor napovedanih varianc  $\boldsymbol{\sigma}_{\boldsymbol{\theta}}^2(\boldsymbol{\rho})$  izraža zaupanje v napovedane srednje vrednosti parametrov v odvisnosti od informacije, ki je na voljo za učenje GP modelov.

### 4.3.2 Simulacija FSGP modela

Simulacija FSGP modela je podobna simulaciji LMN z VBL pristopom v dveh pogledih:

- sistem simuliramo v zveznem prostoru z zveznimi signali in
- namesto signalov  $u(t), y(t)$  uporabljamo njihove časovne odvode  $\dot{u}(t), \dot{y}(t)$ . Z uporabo odvodov signalov namesto signalov se namreč znebimo enosmerne komponente identificiranih afinih lokalnih modelov, v enačbi (4.28) predstavljenih s členom  $(\mathbf{F}_0 - \mathbf{x}_0)$ .

Vrednosti signalov med simulacijo uporabljamo samo, če njihove vzorčene vrednosti nastopajo kot razvrstilne spremenljivke v vektorju  $\boldsymbol{\rho}(t)$ . V tem primeru

jih potrebujemo, ker nastopajo kot vhodi v GP modele, ki jih uporabljamo za napovedovanje vrednosti parametrov lokalnih modelov  $\theta$ .

Za simulacijo FSGP modela od časa  $t = t_0$  naprej potrebujemo vrednosti odvodov signalov  $\hat{u}(t), \hat{y}(t)$  od časa  $t = t_0 - LT$  in vrednosti razvrstilnih spremenljivk od časa  $t = t_0 - T$  naprej.

Napovedani izhod sistema  $\hat{y}$  v času  $t = t_1$  dobimo kot:

$$\hat{y}(t_1) = y(t_0) + \int_{t_0}^{t_1} \hat{y}(t) dt, \quad (4.41)$$

kjer  $\hat{y}(t)$  dobimo iz FSGP modela. Za izračun napovedanega odvoda signala  $\hat{y}$  v času  $t$  potrebujemo vrednosti odvodov signalov  $\mathbf{z}(t - T)$  in napovedi parametrov  $\hat{\theta}(t - T)$ :

$$\hat{y}(t) = \hat{\theta}(t - T) \mathbf{z}(t - T), \quad (4.42)$$

kjer so  $\hat{\theta}(\boldsymbol{\rho}(t - T))$  z GP modeli napovedani parametri FSGP modela in  $\mathbf{z}(t - T)$  vrednosti odvodov signalov:  $\mathbf{z}(t - T) = [\dot{y}(t - T) \dots \dot{y}(t - LT) \dot{u}(t - T) \dots \dot{u}(t - LT)]^T$ . Ta zapis je le druga oblika zapisa (4.32) in (4.33).

Iz enačbe (4.42) lahko razberemo, da pri izračunu napovedanega odvoda izhoda  $\hat{y}(t)$  uporabljamo več sklopov ocen. Prvi sklop so ocene parametrov  $\theta(t - T) = \hat{\theta}(t - T)$ . Drugi sklop so ocenjene vrednosti odvoda izhoda  $[\hat{y}(t - T) \dots \hat{y}(t - LT)]$ , ker pravih za čase  $t > t_0$  ne poznamo. Tretji potencialni sklop ocen so tiste napovedane vrednosti izhoda  $[\hat{y}(t - kT)]$ ,  $k \in \mathbb{N}$ , ki nastopajo kot razvrstilne spremenljivke.

Vse te ocene bi bile v primeru natančnih vrednosti vhodov (to pomeni, da imajo vhodi le eno natančno določljivo vrednost) v obliki Gaussovih porazdelitev. Podobno kot pri simulaciji običajnega GP modela imamo možnosti, da mero zaupanja, dano z variancami napovedi, propagiramo ali pa ne. Najbolj enostavna možnost je, da pri napovedovanju parametrov  $\theta$  predpostavljamo, da vrednosti odvodov signalov  $\mathbf{z}(t - T)$  in vrednosti razvrstilnih spremenljivk  $\boldsymbol{\rho}(t - T)$  niso naključne spremenljivke oz. njihove vrednosti nadomestimo s srednjimi vrednostmi naključnih spremenljivk. Analitična izpeljava izrazov za napovedovanje pri tej predpostavki je v Dodatku B.2.

Izpeljava za propagacijo celotne nedoločenosti skozi FSGP model bi bila precej obsežnejša in nerodna za implementacijo. Alternativa analitični propagaciji negotovosti je propagacija negotovosti z numeričnimi metodami, podobna tisti pri GP modelu. Po tem pristopu v vsakem koraku teka simulacije za vrednost spremenljivk, danih v obliki porazdelitve, uporabimo naključno izbrane vrednosti iz teh porazdelitev. Izvedemo  $S$  tekov simulacije in na koncu rezultate simulacij povprečimo. Izvedba take simulacije ni preveč zahtevna, se pa sorazmerno s številom izvedenih tekov simulacije  $S$  poveča računsko breme.

Če se zadovoljimo z rezultati, ki samo nakazujejo stopnjo zaupanja v rezultate simulacije, nedoločenosti, dane z variancami vseh uporabljenih napovedi, ni treba propagirati. To še posebej velja za primere, pri katerih je pomembno hitro napovedovanje izhoda FSGP modela. Za oceno negotovosti lahko uporabimo varianco parametrov, ki pove, v kako dobro opisanem področju je model. Varianc parametrov je toliko kot parametrov in bi jih lahko združili v en pokazatelj, npr. uteženo vsoto:

$$\text{var}_{\boldsymbol{\theta}} = \frac{1}{2L} \sum_{i=1}^{2L} \varpi_i \sigma_{\theta_i}^2, \quad (4.43)$$

kjer so  $\varpi_i$  uteži posameznih napovedi varianc. Napovedane variance bi lahko utežili enakomerno ali pa relativno glede na pomembnost posameznih regresorjev.

Pri simulaciji FSGP modela smo govorili o zveznih signalih, čeprav danes za skoraj vse simulacije uporabljamo digitalne računalnike. Simulacijo v zveznem prostoru izvedemo z uporabo ustreznih orodij, ki zvezen prostor lahko aproksimirajo, npr. Matlabovo simulacijsko okolje Simulink. Shema simulacije FSGP modela sistema drugega reda z eno razvrstilno spremenljivko v Matlab/Simulinku je prikazana na sliki 4.10.

Naj na tem mestu omenimo še alternativno možnost, to je izpeljava iz razdelka 4.3.1 v diskretnem prostoru. V tem primeru namesto zveznih odvodov uporabimo diskretno aproksimacijo, ki pa se je izkazala za pregrubo in ni omogočala ponazoritve nelinearnega sistema z diskretnimi lokalnimi modeli in uporabo hitrostne linearizacije.

### 4.3.3 Postopek identifikacije FSGP modela

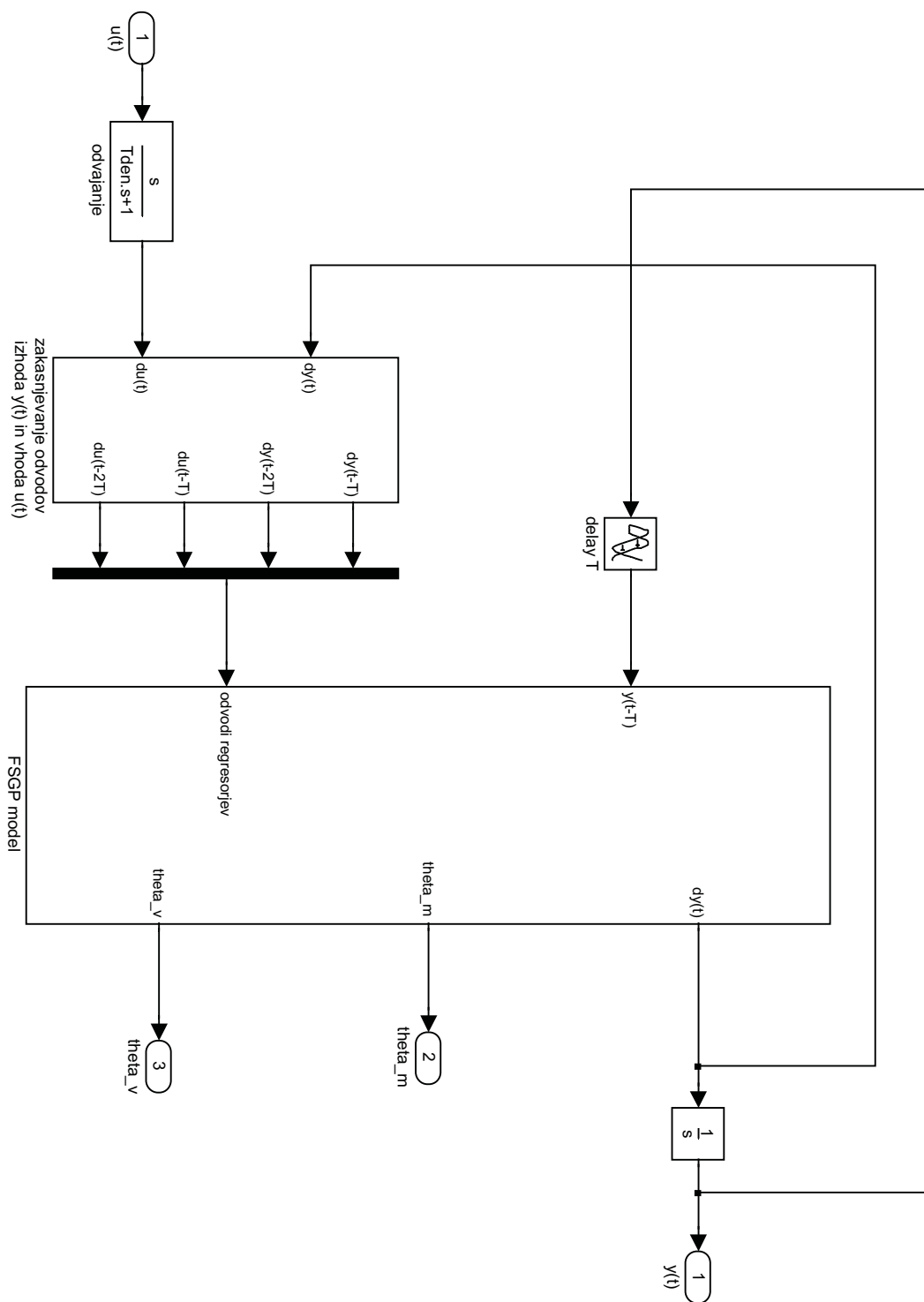
Oglejmo si, kakšen je v grobem postopek identifikacije FSGP modela. Podobno kot pri identifikaciji drugih modelov je tudi ta postopek iterativen. Vrnitev na kakšnega izmed prejšnjih korakov je vedno možna in ponavadi nujna.

#### 1. *Izbira modela*

Odločitev o izbiri modela izhaja (tudi) iz namena uporabe. FSGP model izberemo, kadar želimo za opis neznanega sistema parametrični model. Vzrok za tako izbiro je na primer načrtovanje vodenja, pri čemer parametrični model nudi večjo izbiro načrtovalnih metod.

#### 2. *Razdelitev področja delovanja na podpodročja in eksperiment*

Podatki, iz katerih gradimo FSGP model, so lokalni modeli oz. njihovi parametri. Pred njihovo identifikacijo moramo področje delovanja sistema primerno razdeliti na podpodročja. Pri tem moramo paziti na dvoje:



**Slika 4.10:** Simulacijska shema paketa Matlab/Simulink za FSGP model drugega reda brez propagacije negotovosti. Razvrstilna spremenljivka  $\rho(t)$  je ocena izhoda sistema v prejšnjem koraku  $\hat{y}(t-T)$ . S  $theta_m$  in  $theta_v$  sta označeni napovedana srednja vrednost  $\mu_{\theta}(\rho)$  in varianca  $\sigma_{\theta}^2(\rho)$  parametrov.

- Sistem mora biti dobro opisan na celotnem področju delovanja. Pri določanju števila in lege modelov si lahko pomagamo z metodami, ki so znane iz identifikacije z LMN [32]. V splošnem eksperiment načrtujemo tako, da so z lokalnimi modeli opisana tako ravnotežna kot neravnotežna področja sistema.
- Vpliv lokalnih modelov je določen z vrednostmi razvrstilnih spremenljivk. Ker vrednosti razvrstilnih spremenljivk nastopajo kot učni vhodi v GP modele, je potrebno paziti, da te vrednosti nosijo zadosti informacije za učenje GP modelov.

Razvrstilne spremenljivke so navadno neka podmnožica regresorjev lokalnih modelov. Če v vektorju razvrstilnih spremenljivk nastopa več regresorjev iz istega signala, so njihove vrednosti v ravnotežnih točkah sistema enake; to pomeni, da je za optimizacijsko metodo prispevek regresorjev iz istega signala enak. To predstavlja razmeroma velik problem, ki ga lahko rešimo tako, da:

- za učenje uporabimo tudi lokalne modele iz neravnotežnega področja ali
- zmanjšamo vektor razvrstilnih spremenljivk.

Prva možnost je privlačnejša, ker tako dobimo boljši opis dinamičnega sistema. Vendar je, kot vemo, identifikacija lokalnih modelov v neravnotežnem področju navadno znatno težja od identifikacije lokalnih modelov na ravnotežnem področju (podpoglavje 4.2). Druga možnost pa lahko zaradi zmanjšanja vektorja razvrstilnih spremenljivk  $\rho$  pripelje do, kot smo omenili že v podpoglavju 3.1, problemov pri zlivanju lokalnih modelov na področjih, oddaljenih od ravnotežne krivulje. Obstaja možnost, da v vektorju razvrstilnih spremenljivk ne nastopajo regresorji lokalnih modelov ali pa nastopa od vsakega signala samo po en regresor. V tem primeru je problem vsebovane informacije učnih vhodov precej manjši.

### 3. Identifikacija lokalnih modelov

Iz izmerjenih podatkov identificiramo lokalne modele z eno izmed običajnih metod za identifikacijo linearnih dinamičnih sistemov [64, 69, 74]. Preden parametre uporabimo za učenje GP modelov, moramo identificirane lokalne modele ovrednotiti.

### 4. Postavitev in učenje GP modelov

Naslednji korak je postavitev in učenje GP modelov, ki opisujejo parametre lokalnih modelov. Kovariančno funkcijo za GP modele kot običajno izberemo glede na predznanje o lastnostih opisovanega sistema, v našem primeru so to lastnosti funkcij, ki opisujejo odvisnost parametrov lokalnih modelov od vrednosti razvrstilnih spremenljivk. Kot izhodni del uče

množice nastopajo vrednosti ustreznega parametra v različnih področjih sistema, kot vhodni del pa ustrezne vrednosti razvrstilnih spremenljivk, tj. tistih regresorjev, ki določajo delovno področje sistema. Z naraščanjem reda modela se hitro povečuje tudi število potrebnih učnih točk oz. lokalnih modelov (prekletstvo dimenzije). Problem lahko zmanjšamo z zmanjšanjem vektorja razvrstilnih spremenljivk, ki pa ima, kot je omenjeno v točki 2, svoje probleme. Pri določanju potrebnih regresorjev, tj. zakasnjenih vhodov in izhodov, ki vplivajo na posamezne parametre, si lahko pomagamo z uporabo ARD lastnosti GP modela ali s predznanjem o sistemu.

#### 5. FSGP model in njegovo vrednotenje

Naučeni GP modeli sestavljajo FSGP model, opisan z enačbo (4.40). Pred njegovo uporabo je potrebno še njegovo ovrednotenje, kot je opisano v podglavju 3.6.

### 4.3.4 Primer identifikacije dinamičnega sistema drugega reda s FSGP modelom

Identifikacijo dinamičnih sistemov s FSGP modelom bomo predstavili na nelinearnem dinamičnem sistemu (4.2). Začnemo z zapisom njegovega zveznega ekvivalenta:

$$\begin{aligned} y(t+T) &= f(y(t), y(t-T), u(t)) = \\ &= k_1 y(t) + k_2 y^2(t) + k_3 y(t-T) + k_3 u(t)y(t) + k_4 u(t) \end{aligned} \quad (4.44)$$

s konstantami  $k_1 = 0.893$ ,  $k_2 = 0.0371$ ,  $k_3 = -0.05$  in  $k_4 = 0.157$  in časom zakasnitve  $T = 1$  s. Delovno področje sistema smo omejili z vhodnim signalom  $1 < u(t) < 3$ , kar ustreza vrednostim izhoda  $y(t)$  v ravnotežju  $0.91 < y(t) < 2.03$ .

Sistem (4.44) bomo identificirali s FSGP modelom drugega reda:

$$\dot{y}(t+T) = -a_1(\boldsymbol{\rho}) \dot{y}(t) - a_2(\boldsymbol{\rho}) \dot{y}(t-T) + b_1(\boldsymbol{\rho}) \dot{u}(t) - b_2(\boldsymbol{\rho}) \dot{u}(t-T) \quad (4.45)$$

s spremenljivimi parametri  $-a_1(\boldsymbol{\rho}) = \frac{\partial y(t+T)}{\partial y(t)}$ ,  $-a_2(\boldsymbol{\rho}) = \frac{\partial y(t+T)}{\partial y(t-T)}$ ,  $b_1(\boldsymbol{\rho}) = \frac{\partial y(t+T)}{\partial u(t)}$  ter  $b_2(\boldsymbol{\rho}) = \frac{\partial y(t+T)}{\partial u(t-T)}$ . Odvisnost parametrov lokalnih modelov  $a_1$ ,  $a_2$ ,  $b_1$  in  $b_2$  od območja delovanja sistema bomo poskusili najprej opisati z vektorjem regresorjev:  $\boldsymbol{\rho}(t) = [u(t)y(t)]^T$ .

Za učenje GP modelov, ki sestavljajo FSGP model, potrebujemo parametre lokalnih modelov, zato identificiramo lokalne modele drugega reda oblike:

$$y(t+T) = -a_1 y(t) - a_2 y(t-T) + b_1 u(t) - b_2 u(t-T) \quad (4.46)$$

na zelenem delovnem področju tako na ravnotežni krivulji, kot na področjih stran od nje. Za identifikacijo lokalnih modelov na ravnotežni krivulji smo sistem

najprej pripeljali v delovno točko, določeno z  $U_0, Y_0$ , nato pa vhodnemu signalu dodali PRBS signal z amplitudo  $\Delta u = 0.025$ . Parametre lokalnih modelov  $a_1, a_2, b_1$  in  $b_2$  iz signalov  $u(t), y(t)$  poiščemo z izbrano metodo za identifikacijo linearnih modelov.

Na področjih stran od ravnotežne krivulje je identifikacija lokalnih modelov navadno težja, saj so ta področja opisana z manj podatki. Potrebujemo vhodni signal, ki sistem dovolj vzbuja, da dobimo dovolj informacije na posameznih podpodročjih, na katerih želimo identificirati lokalne modele. Zato smo vhodni signal  $u(t)$  načrtovali v dveh delih. Najprej smo naredili njegovo osnovo, tj. signal  $u_o(t)$  dolžine 254 s, ki je naključno spremenil svojo vrednost vsake  $T_{sw} = 2$  s. Iz tega signala smo nato naredili  $N_{\text{sig}} = 300$  različnih signalov, tako, da smo vrednostim signala  $u_o(t)$  dodali beli Gaussov šum s standardno deviacijo  $\sigma_u = 0.01$ . S tako realiziranimi signali dosežemo, da se sistem premika vedno po skoraj enaki poti, odmiki od te poti pa nastanejo zaradi vhodu dodanega šuma.

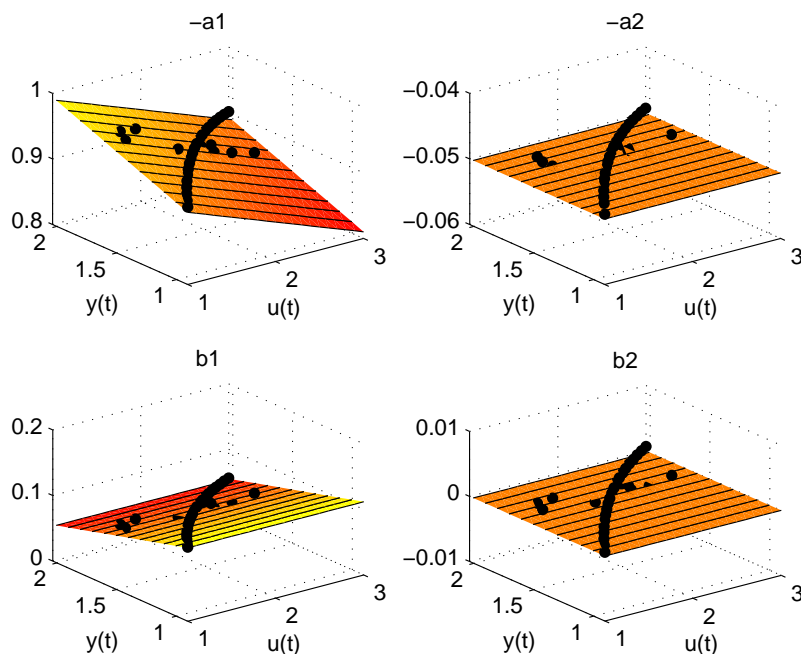
Za identifikacijo posameznega lokalnega modela:

$$y(kT + T) = -a_1y(kT) - a_2y(kT - T) + b_1u(kT) - b_2u(kT - T) \quad (4.47)$$

uporabimo vzorčene vrednosti iz vseh tako dobljenih signalov z istim odmikom  $kT$ , saj je okolica točke, določene z  $[y(kT) \ y(kT - T) \ u(kT) \ u(kT - T)]$ , opisana z  $N_{\text{sig}} = 300$  vzorci. Lokalne modele smo identificirali z metodo najmanjših kvadratov s posplošenim pogreškom.

Na ravnotežni krivulji smo identificirali dvajset modelov, na področjih, ki so od nje odmaknjena, pa devet. Vrednosti identificiranih parametrov v odvisnosti od izbrane razvrstilne spremenljivke  $\boldsymbol{\rho}(t) = [u(t)y(t)]^T$  so, skupaj s pravimi vrednostmi, prikazane na sliki 4.11.

Identificirana parametra  $-a_2$  in  $b_2$  sta konstantna, zato ju nismo opisali z GP modeloma, ampak smo jima določili kar konstantne vrednosti:  $-a_2 = -0.05$  in  $b_2 = 0$ . Parametra  $a_1$  in  $b_1$  smo opisali vsakega s svojim GP modelom, pri čemer smo za vhod v GP modela uporabili vrednosti regresorjev  $[x(t - T)u(t - T)]$ . Uporabili smo Gaussovo kovariančno funkcijo (2.4), hiperparametre GP modelov pa smo optimirali z ML metodo. Vrednosti hiperparametrov za naučena GP modela so zbrane v tabeli 4.2. Iz druge vrstice tabele vidimo, da je vrednost hiperparametra  $w_2$  GP modela za parameter  $b_1$  zanemarljiva, zato smo regresor  $u(t)$  izločili iz vektorja razvrstilnih spremenljivk  $\boldsymbol{\rho}(t)$  za GP model, ki opisuje parameter  $b_1$ , in modelu še enkrat optimirali hiperparametre. Rezultati napovedovanja obeh GP modelov so skupaj z analitično dobljenimi vrednostmi



**Slika 4.11:** Vrednosti identificiranih parametrov lokalnih modelov v odvisnosti od regresorjev  $u(t)$  in  $y(t)$ . Površine na slikah predstavljajo prave, analitično izračunane, vrednosti parametrov  $-a_1$ ,  $-a_2$ ,  $b_1$  in  $b_2$ .

za primerjavo

$$-a_1(\boldsymbol{\rho}(t)) = \frac{\partial y(t+T)}{\partial y(t)} = k_1 + 2k_2 y(t) + k_3 u(t), \quad (4.48)$$

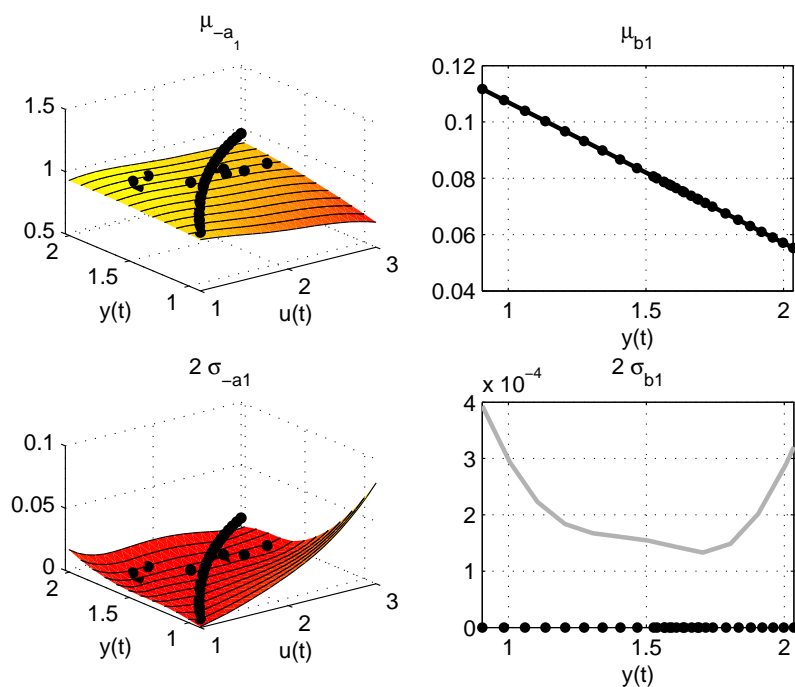
$$-a_2(\boldsymbol{\rho}(t)) = \frac{\partial y(t+T)}{\partial y(t-T)} = k_3, \quad (4.49)$$

$$b_1(\boldsymbol{\rho}(t)) = \frac{\partial y(t+T)}{\partial u(t)} = k_3 y(t) + k_4, \quad (4.50)$$

$$b_2(\boldsymbol{\rho}(t)) = \frac{\partial y(t+T)}{\partial u(t-T)} = 0 \quad (4.51)$$

prikazani na sliki 4.12.

Tako dobljen FSGP model je bil ovrednoten s simulacijo. Za napovedovanje parametrov, odvisnih od razvrstilnih spremenljivk, smo uporabili naučena GP modela. Odziva sistema in modela na naključni vhodni signal sta prikazana na sliki 4.13 zgoraj. Na začetku, ko je sistem na dobro opisanem področju delovanja, FSGP model lepo posnema obnašanje sistema. Ko pa se sistem zaradi spremembe vhoda premakne v področje, ki z modelom ni opisano, se napovedovanje modela poslabša, vendar nas na to opozori povečana mera negotovosti za napovedana parametra  $a_1$  in  $b_2$ , kar je lepo vidno proti koncu simulacije.



**Slika 4.12:** Napovedana srednja vrednost  $\mu_{-a_1}$  in mera negotovosti  $2\sigma_{-a_1}$  parametra  $a_1$  (levo) ter napovedana srednja vrednost  $\mu_{b_1}$  in mera negotovosti  $2\sigma_{b_1}$  parametra  $b_1$  (desno)

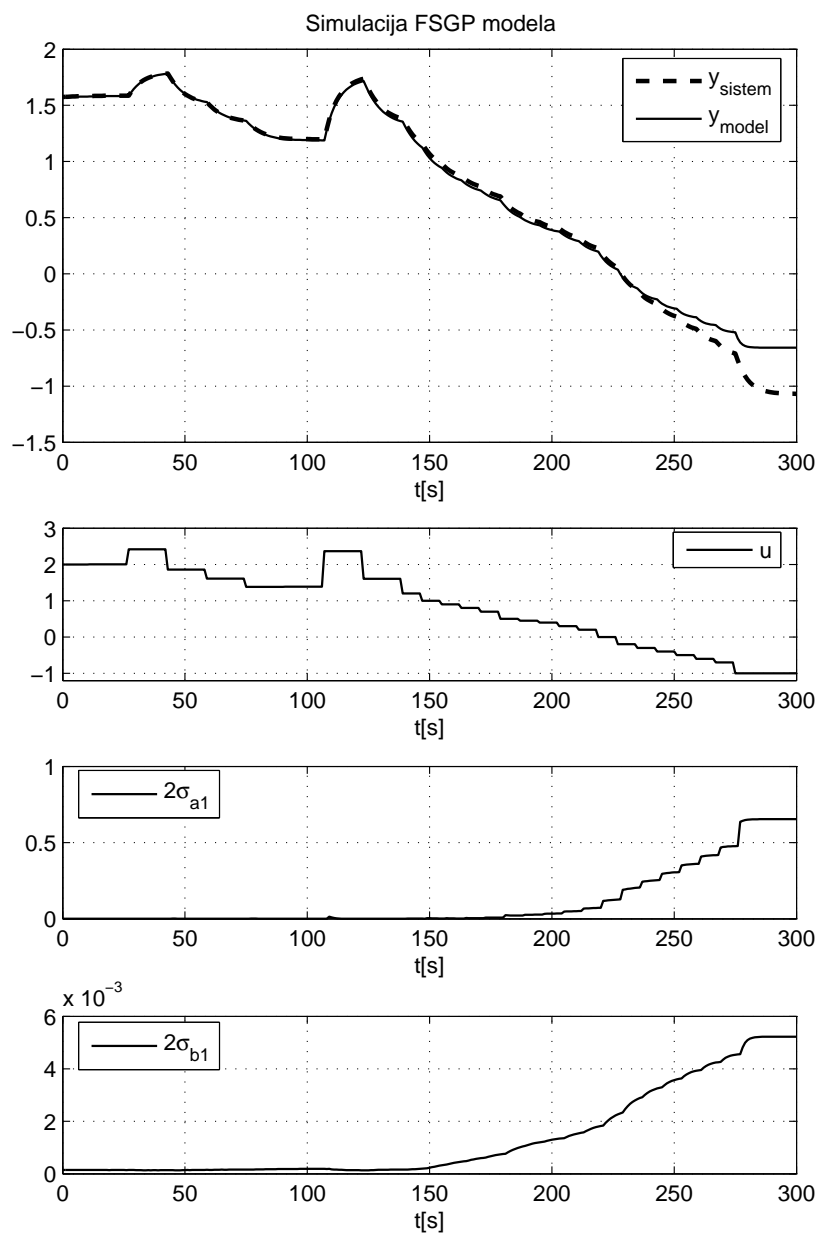
**Tabela 4.2:** Vrednosti hiperparametrov GP modelov za parameter  $a_1$  in  $b_1$

	$w_1$	$w_2$	$v$	$v_0$
GP $_{a_1}$	0.18	0.49	0.39	$1 \cdot 10^{-12}^\dagger$
GP $_{b_1}$	$7 \cdot 10^{-3}$	$8 \cdot 10^{-13}$	0.13	$1 \cdot 10^{-13}^\dagger$
♣GP $_{b_1}$	$7 \cdot 10^{-3}$	×	0.13	$1 \cdot 10^{-13}^\dagger$

Opombe:

† V identifikacijskih signalih ni bilo prisotnega šuma, ocena variance šuma  $v_0$  je zelo nizka;

♣ model z odstranjenim regresorjem  $u(t)$ .



**Slika 4.13:** Primerjava odziva sistema in FSGP modela na dobro opisanem področju ( $0 < t < 150$  s) in slabo opisanem področju ( $150 < t < 300$  s). Na drugi sliki vhodni signal  $u(t)$ , na spodnjih dveh mere negotovosti za napovedana parametra FSGP modela  $\sigma_{a1}(t)$  in  $\sigma_{b1}(t)$

\* \* \*

Glavna lastnost, po kateri se FSGP model loči od drugih izpeljank GP modela, je njegova parametrična struktura. Ta je prednost modela pri nekaterih oblikah uporabe, npr. pri načrtovanju vodenja.

Razlika med FSGP modelom in običajno mrežo lokalnih modelov, realizirano z VBL pristopom, je zlivanje parametrov lokalnih modelov. V LMN so za zlivanje uporabljene utežnostne funkcije, medtem ko so v FSGP modelu za zlivanje parametrov uporabljeni GP modeli. Slednja izvedba zlivanja ima lahko nekaj potencialnih prednosti:

1. Napovedovanje z uporabo GP modela daje dobre rezultate pri reševanju regresijskih problemov z majhnim številom točk, med katerimi moramo interpolirati.
2. Izhod GP modelov niso samo najbolj verjetne vrednosti parametrov, ampak tudi odgovarjajoče mere zaupanja v te napovedi. Te dajo dodatno informacijo o kvaliteti modela, ki jo lahko koristno uporabimo, npr. v vodenju, pri katerem z opazovanjem napovedanih varianc parametrov sistem lahko obdržimo v boljše opisanih področjih, tj. področjih z manjšo varianco, kot smo pokazali v [9].
3. Pri izbiri razvrstilnih spremenljivk si lahko pomagamo z ARD lastnostjo GP modelov, kot smo opisali v tretjem poglavju in pokazali na primeru.
4. GP modeli ponujajo možnost vključevanja informacije o kvaliteti identificiranih parametrov lokalnih modelov v FSGP model.

Pri identifikaciji dinamičnih sistemov s FSGP modelom se lahko pojavi tudi nekaj problemov. Če želimo dober opis celotnega sistema, morajo posamezni lokalni modeli:

- dovolj dobro opisovati svoje področje in
- biti na področjih, kjer nas zanima delovanje sistema. Če nas zanima model tudi v oddaljenosti od ravnotežnega področja, potrebujemo informacijo, tj. lokalne modele, tudi na tem področju. Identifikacija lokalnih modelov na neravnotežnih področjih je navadno otežena zaradi manjšega števila podatkov, ki ta področja opisujejo.

Opisani problem lahko zmanjšamo z zmanjšanjem vektorja razvrstilnih spremenljivk. Krajši (manj popolni) vektor razvrstilnih spremenljivk izberemo, manj podatkov potrebujemo (prekletstvo dimenzije), toda hkrati slabše sklepamo na odvisnost parametrov lokalnih modelov od delovnega področja. Uporaba zmanjšane vektorja razvrstilnih spremenljivk ima lahko za posledico nezvezen in celo nestabilen FSGP model. Stabilnost posameznih lokalnih modelov je potreben, ni pa zadosten pogoj za stabilnost dobljenega FSGP modela. Opisane problematične lastnosti so sicer skupne vsem pristopom, ki temeljijo na mreži lokalnih modelov.

FSGP modela ne smemo zamenjati z linearnim GP modelom (angl. *Linear GP model*) [32], pri katerem gradimo mrežo lokalnih modelov (LMN) v kateri je vsak posamezni lokalni model predstavljen z GP modelom.

#### 4.4 Ilustracija drugih načinov vključevanja predznanja v GP model

V tem podpoglavju bomo predstavili preostale oblike vključevanja predznanja v GP model:

- vključevanje statične karakteristike,
- vključevanje znanja o histerezi in
- vključevanje znanja o šumnem procesu na izhodu sistema, pri čemer bomo predpostavili barvni Gaussov šum.

Uporabnost vključevanja predznanja bomo prikazali s primerjavami dveh GP modelov – enega z vključenim in drugega brez vključenega predznanja. Modela primerjamo na podlagi vrednotenja rezultatov simulacije, kot je bilo opisano v podpoglavju 3.6. Zanimala nas bo:

- kvalitativna primerjava rezultatov simulacije obeh modelov in
- kvantitativna primerjava rezultatov simulacije obeh modelov, za katero bomo uporabljali cenilke: povprečna kvadratna napaka SE (3.20), povprečna relativna kvadratna napaka MRSE (3.21) in logaritem gostote napake LD (3.22).

#### 4.4.1 Vključevanje predznanja o statični karakteristiki

Znano statično karakteristiko sistema lahko vključimo v učne podatke  $\mathcal{D} = \mathbf{X}, \mathbf{y}$  z dodajanjem učnih točk, ki jo opisujejo.

Za ravnotežne točke dinamičnega sistema velja  $u(k-1) = \dots = u(k-L) = u_{eq}$  in  $y(k) = y(k-1) = \dots = y(k-L) = y_{eq}$ , kjer je  $L$  red sistema,  $y(k)$  in  $u(k)$  pa so vzorčene vrednosti vhoda v sistem  $u$  in izhoda iz sistema  $y$  v času  $t = kT$ .

Pri opisu sistema  $L$ -tega reda v vhodnem vektorju  $\mathbf{x}_i$  navadno nastopajo regresorji  $\mathbf{x}_i = [y(k-1) \dots y(k-L) \ u(k-1) \dots u(k-L)]$ , razdelek 3.5, učni izhod pa predstavlja  $y_i = y(k)$ . Posamično točko statične karakteristike tako lahko dodamo kot učni par  $\mathbf{x}_i, y_i$ , pri čemer je vhodni del sestavljen kot  $\mathbf{x}_i = [\mathbf{y}_{eq} \ \mathbf{u}_{eq}]$ , izhodni pa kot  $y_i = y_{eq}$  in sta vektorja  $\mathbf{y}_{eq}$  in  $\mathbf{u}_{eq}$  vektorja z  $L$  elementi  $y_{eq}$  ter  $u_{eq}$ .

Vključevanje predznanja o statični karakteristiki je bilo na primeru pokazano skupaj z vključevanjem linearnih lokalnih modelov (podpoglavje 4.2). Centri lokalnih modelov na ravnotežni krivulji so namreč hkrati tudi točke ravnotežne krivulje, tj. statične karakteristike.

#### 4.4.2 Vključevanje znanja o histerezi v GP model

Histereza sistema je pojav, pri katerem odziv sistema na vzbujanje ni odvisen samo od velikosti vzbujanja, ampak tudi od "zgodovine" sistema. Odziv sistema na vzbujanje lahko opišemo z dvema funkcijama; ena funkcija opisuje odziv sistema na vzbujanje v eni smeri, druga funkcija pa opisuje odziv sistema na vzbujanje v nasprotni smeri [18]. Drugače: izhod sistema  $y$  ni odvisen samo od njegove (bližnje) zgodovine  $\mathbf{x}_k$  (3.2), temveč tudi od stanja sistema  $\eta(k)$ :

$$y(k) = f(\mathbf{x}_k, \eta(k)) + \epsilon(k), \quad (4.52)$$

kjer je  $\epsilon(k)$  šum na izhodu sistema. Je posebna oblika nelinearnosti, saj se sistem s histerezo obnaša, kot bi imel "spomin".

Ena izmed možnosti, kako v GP model vključiti znanje o histerezi, je, kot smo pokazali v [8], razširitev vhodnega vektorja  $\mathbf{x}_k$  (3.2) še za en element, ki predstavlja stanje sistema  $\eta$ . Razširjeni vektor regresorjev  $\mathbf{x}_k$  je tako:

$$\mathbf{x}_k = [y(k-1) \dots y(k-L) \ u(k-1) \dots u(k-L) \ \eta(k)]. \quad (4.53)$$

Pri sestavljanju učnih podatkov  $\mathcal{D} = \mathbf{X}, \mathbf{y}$  vsakemu vhodnemu vzorcu  $\mathbf{x}_i$  dodamo še stanje histereze  $\eta(i)$  v koraku  $i$ . Enako postopamo tudi pri točkah, ki predstavljajo (morebitno) vključeno statično karakteristiko.

## Primer

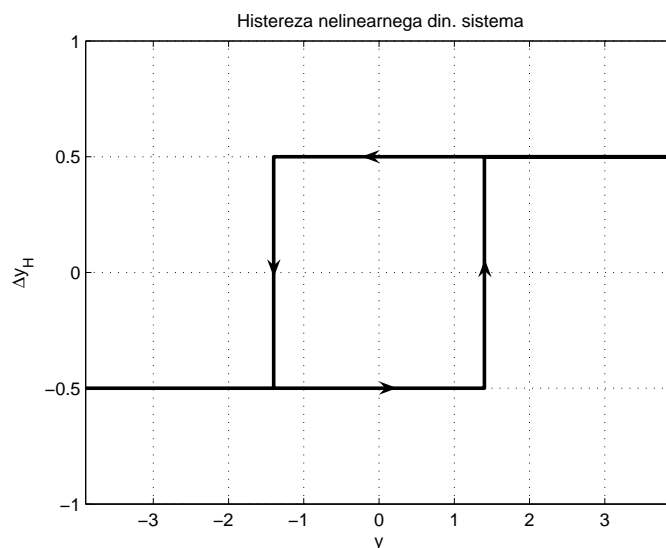
Vključevanje histereze bomo predstavili na primeru nelinearnega dinamičnega sistema. Za osnovo bomo vzeli sistem (4.1), ki ima na izhodu histerezo z  $\Delta y_H = \pm 0.5$  in širino  $-1.4 < y < 1.4$  (slika 4.14). Za izhod iz sistema velja:

$$y_H(k+1) = y(k+1) + \Delta y_H \eta(k) + \epsilon(k). \quad (4.54)$$

V vsakem koraku  $k$  se po določitvi izhoda  $y(k+1)$  (slika 4.15) posodobi tudi stanje sistema (histereze)  $\eta(k+1)$ :

$$\eta(k+1) = \begin{cases} 1, & \text{ko } y(k) \geq l_h = 1.4 \\ -1, & \text{ko } y(k) \leq l_l = -1.4 \\ \eta(k), & \text{drugače.} \end{cases} \quad (4.55)$$

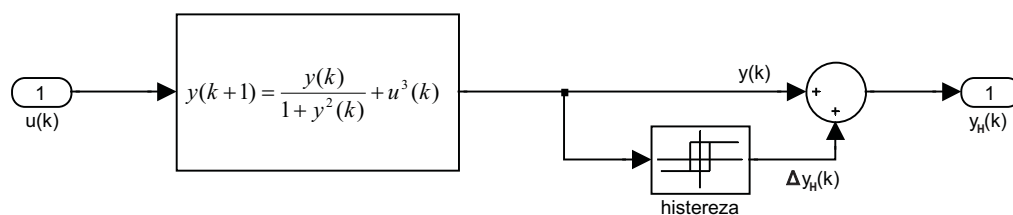
Izhod sistema  $y_H(k+1)$  je moten z belim šumom  $\epsilon(k)$  z varianco  $v_0 = 4 \cdot 10^{-4}$ . Statična karakteristika sistema je prikazana na sliki 4.16.



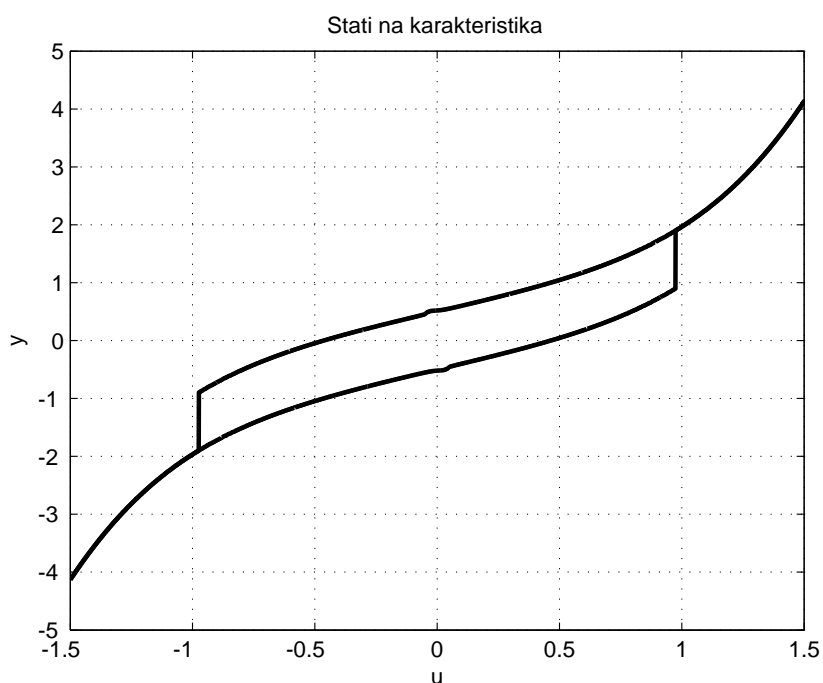
**Slika 4.14:** Histereza sistema, ki ga želimo identificirati

Naš cilj je bilo zgraditi GP model, ki bo opisoval dinamično obnašanje sistema (4.54) na področju, omejenem z vhom  $u$  med  $u_{\min} = -1.5$  in  $u_{\max} = 1.5$ . Za ovrednotenje dobljenega modela smo primerjali rezultate simulacije GP modela z obnašanjem sistema. Uporabili smo dva različna GP modela z Gaussovo kovariančno funkcijo (2.4):

- GP model brez predznanja o histerezi in



**Slika 4.15:** Shema sistema, ki ga želimo identificirati, v programskem paketu Matlab/Simulink



**Slika 4.16:** Statična karakteristika sistema s histerezo

- GP model z vključenim predznanjem o histerezi.

Za učenje smo uporabili vzorce iz statične karakteristike sistema in vzorce, ki smo jih dobili iz dinamičnega odziva modela. Ko nas je zanimala dinamika sistema, smo za njegovo vzbujanje uporabili diskretni psevdo-naključni signal (str. 74) s taktom  $T_{sw} = 3$ , katerega amplituda ob spremembi lahko zavzame katero koli vrednost med  $u_{\min}$  in  $u_{\max}$ .

Najprej je bil narejen GP model za sistem (4.54) brez vključevanja znanja o histerezi v model. Iz statične karakteristike in dinamičnega odziva modela smo z vzorčenjem dobili 242 parov vhodno/izhodnih podatkov, ki so opisovali obnašanje sistema in so bili uporabljeni za učenje. Ob vzorčenju v  $k$ -tem koraku sta vhodni del učnega vektorja sestavljala vhod in izhod iz sistema  $\mathbf{x}_k = [u(k-1), y(k-1)]$ , izhodni del učnega vektorja pa izhod  $y(k)$ .

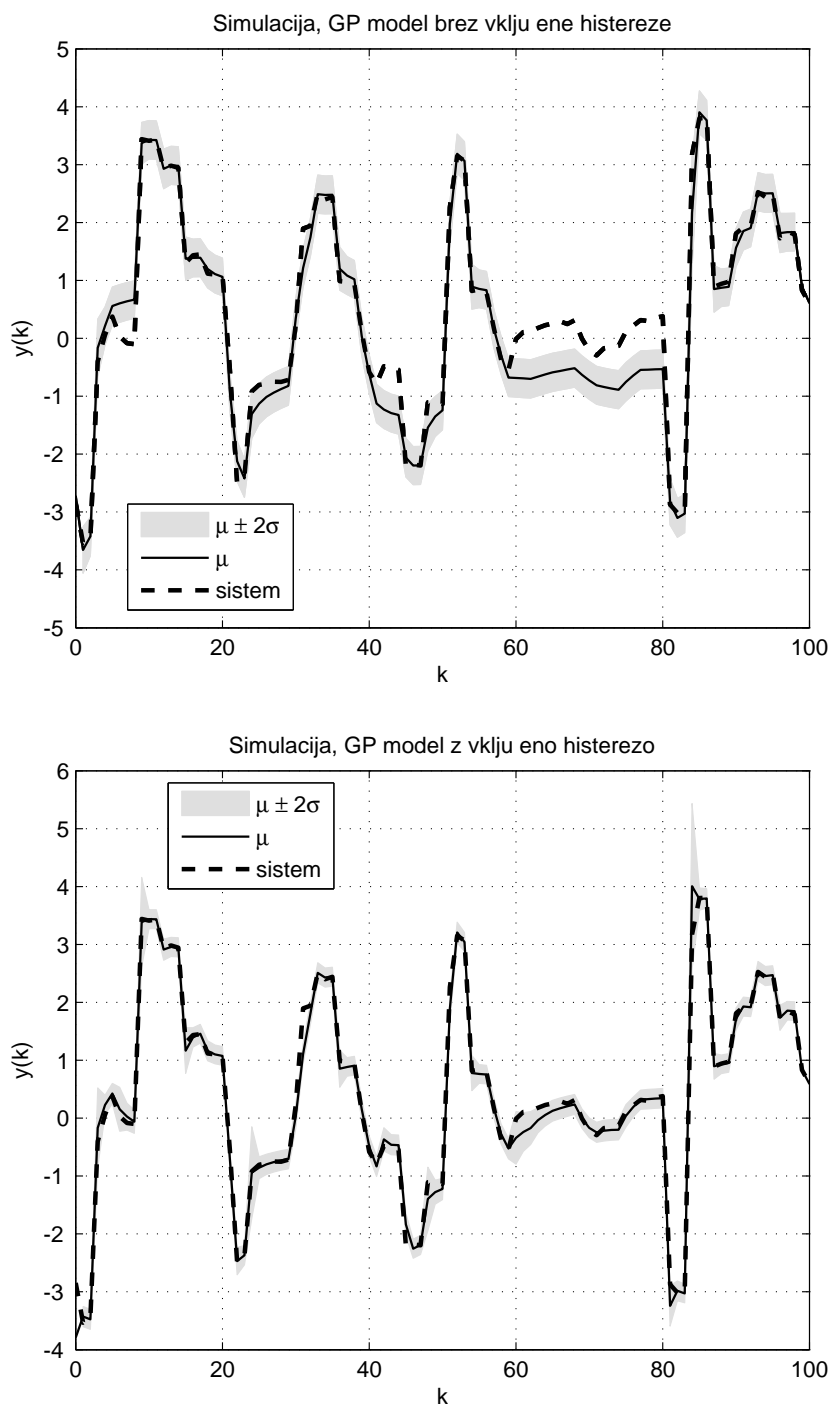
Za učenje drugega modela, tj. GP modela z vključenim znanjem o histerezi, so bili uporabljeni isti vzorci kot za učenje prvega modela, le da so imeli vhodni deli učnih vzorcev eno komponento več – podatek o stanju sistema  $\eta(k)$ :  $\mathbf{x}_k = [u(k-1), y(k-1), \eta(k)]$ .

Za primerjavo smo oba modela vzbujali z drugim, a podobnim signalom, kot je bil uporabljen za učenje, in primerjali rezultate simulacije s pravim odzivom modela. Zaradi ilustrativosti so prikazani rezultati le za manjši del tega signala. Rezultate ovrednotenja obeh modelov vidimo na sliki 4.17.

Opazno je boljše obnašanje drugega modela, saj ima prvi model nekaj težav na področju  $l_l \leq y(k) \leq l_h$ , kjer ima sistem histerezo obnašanje. Pričakovali bi, da bo prvi model, glede na to da ne pozna histereze, opozoril na večjo negotovost napovedi s povečanjem variance. Zakaj se to ne zgodi v večjem obsegu? Manjkajoča informacija o histerezi, v drugem modelu izražena z regresorjem  $\eta$ , se izrazi prek hiperparametra  $v_0$ , ki izraža varianco šuma. Ta z učenjem dobi vrednost  $v_0 = 2.6 \cdot 10^{-2}$  za model brez vključenega predznanja in vrednost  $v_0 = 6.0 \cdot 10^{-3}$  za model z vključenim predznanjem. Z izbrano kovariančno funkcijo predpostavimo enak šum na celotnem področju delovanja, vendar mora prvi GP model v to zajeti še “šum” zaradi histereze. Tako je hiperparameter  $v_0$  prvega modela ocenjen previsoko glede na varianco šuma na izhodu, a še vedno prenizko, da bi lahko zajel vse napake v napovedovanju zaradi histereze. Za boljši model potrebujemo več predznanja o strukturi identificiranega sistema. Vrednosti cenilk, zbranih v tabeli 4.3, potrjujejo boljše obnašanje modela z vključenim predznanjem.

**Tabela 4.3:** Vrednosti cenilk za rezultata simulacije GP modelov z vključeno histerezo in brez nje:

Cenilka	“Brez histereze”	“S histerezo”
SE	0.19	0.035
MRSE	0.25	0.11
LD	2.4	-0.41



**Slika 4.17:** Simulacija GP modela brez znanja o histerezi (zgoraj) in z vključenim znanjem o histerezi (spodaj)

### 4.4.3 Poznavanje šumnega procesa na izhodu – vključevanje barvnega šuma v kovariančno matriko

Do sedaj je bil na izhodu iz nelinearnega dinamičnega sistema, ki smo ga želeli opisati, predpostavljen beli Gaussov šum. Če je Gaussov šum barvni in poznamo njegov model, lahko to predznanje vključimo v kovariančno matriko, kot je bilo za statične sisteme pokazano v [73], za dinamične sisteme pa smo prikazali v [11].

V podpoglavju 3.4.1, kjer smo predstavili izbiro kovariančne funkcije, smo razbili kovariančno funkcijo na funkcijski in šumni del (3.3). Pri belem Gaussovem šumu na izhodu neznanega procesa je vrednost šumnega dela kovariance  $C_n(\mathbf{x}_i, \mathbf{x}_j) = v_0\delta_{ij}$  različna od nič samo v primeru, če gre za isti vhodni vektor  $\mathbf{x}_i = \mathbf{x}_j$ . Pri barvnem šumu na izhodu pa so vrednosti šuma korelirane s svojimi preteklimi vrednostmi in kovarianca je različna od nič tudi za različne vhodne vektorje.

Predpostavimo, da je šum na izhodu iz procesa opisan z avtoregresijskim (AR) modelom z drsečim povprečjem (MA)  $r$ -tega reda:

$$e(k) = \epsilon(k) - a_1e(k-1) \cdots - a_re(k-r) + b_1\epsilon(k-1) + \cdots + b_r\epsilon(k-r), \quad (4.56)$$

kjer je  $\epsilon(k)$  beli Gaussov šum. Označimo šumni del kovariance med dvema vhodnima vektorjema  $C_n(e(k), e(k+\tau))$  kot  $C(\tau)$ <sup>7</sup>. Če vstavimo vhodna vektorja  $e_k$  in  $e_{k+\tau}$ :

$$\begin{aligned} e_k &= \epsilon_k + b_1\epsilon_{k-1} + \cdots + b_r\epsilon_{k-r} - a_1e_{k-1} - \cdots - a_re_{k-r} \\ e_{k+\tau} &= \epsilon_{k+\tau} + \cdots + b_r\epsilon_{k+\tau-r} - a_1e_{k+\tau-1} - \cdots - a_re_{k+\tau-r}, \end{aligned}$$

kjer smo časovno odvisnost spremenljivke  $x$  zaradi poenostavitve namesto z  $x(k)$  označili z indeksom  $x_t$ , lahko kovarianco  $C(0)$  zapišemo:

$$\begin{aligned} C(0) &= C(e_k, e_k) = \\ &= \sigma^2 (1 + b_1^2 + \cdots + b_r^2) + C(e_k, -a_1e_{k-1} - \cdots - a_re_{k-r}). \end{aligned} \quad (4.57)$$

Prvi del enačbe (4.57) predstavlja prispevek MA dela in ga zapišemo krajše kot  $B_0$ , drugi del pa predstavlja prispevek AR dela.  $C(0)$  zaradi enakosti  $C(\tau) = C(-\tau)$  lahko zapišemo kot:

$$C(0) = B_0 - a_1C(1) - a_2C(2) - \cdots - a_rC(r). \quad (4.58)$$

Podobno kot zgoraj za  $\tau = 0$  lahko za  $\tau \neq 0$  šumni del kovariance med dvema različnima vhodnima vektorjema  $e_k$  in  $e_{k+\tau}$  zapišemo kot:

$$C(\tau) = B_\tau - a_1C(\tau-1) - \cdots - a_rC(\tau-r), \quad (4.59)$$

<sup>7</sup>V nadaljevanju tega razdelka bomo indeks  $n$ , ki označuje šumni del kovariančne matrike, izpuščali zaradi poenostavitve zapisa.

kjer

$$\begin{aligned} B_\tau &= \sigma^2(1 \cdot b_i + b_1 \cdot b_{i+1} + \dots + b_{r-i} \cdot b_r), \tau \leq r \\ B_\tau &= 0, \tau > r. \end{aligned} \quad (4.60)$$

Iz prvih  $r + 1$  enačb za kovariance  $C(0) \dots C(r)$  dobimo *Yule-Walkerjev* sistem enačb:

$$\begin{bmatrix} 1 & a_1 & \dots & a_{r-1} & a_r \\ a_1 & 1 + a_2 & \dots & a_r & 0 \\ a_2 & a_1 + a_3 & \dots & 0 & 0 \\ \vdots & & & \ddots & \vdots \\ a_r & a_{r-1} & \dots & a_1 & 1 \end{bmatrix} \begin{bmatrix} C(0) \\ C(1) \\ C(2) \\ \vdots \\ C(r) \end{bmatrix} = \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ \vdots \\ B_r \end{bmatrix}. \quad (4.61)$$

Z njegovo rešitvijo dobimo prvih  $r + 1$  elementov šumnega dela kovariančne matrike:  $C(0), \dots, C(r)$ . Preostale elemente šumnega dela kovariančne matrike  $C(\tau)$ ,  $\tau = r + 1, \dots, N - 1$ , dobimo iterativno z rešitvijo enačbe:

$$C(i) = -a_1 C(i - 1) - a_2 C(i - 2) - \dots - a_r C(i - r). \quad (4.62)$$

V kovariančno matriko  $\mathbf{K}$  vstavimo kovariance med posameznimi vhodnimi vektorji  $\mathbf{x}_i$  in  $\mathbf{x}_j$  kot:

$$C(\mathbf{x}_i, \mathbf{x}_j) = C_f(\mathbf{x}_i, \mathbf{x}_j) + C_n(|i - j|). \quad (4.63)$$

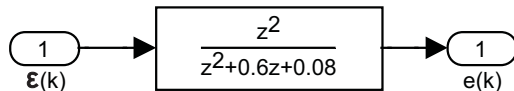
Psevdo-koda opisanega postopka je v Dodatku C.2.

V opisanem postopku smo predpostavili predznanje o modelu šuma. Če tega znanja nimamo, lahko skušamo model šuma pridobiti iterativno z optimizacijo celotnega modela procesa, kot je običajno pri identifikaciji parametričnih modelov.

## Primer

Tudi za prikaz vključevanja predznanja o šumu smo uporabili sistem (4.1), le da smo na izhodu iz sistema predpostavili barvni šum  $e(k)$  z varianco  $v_0 = 0.01$ , ki ga lahko opišemo z ARMA modelom s parametri  $a_1 = -0.6$ ,  $a_2 = 0.08$ ,  $b_0 = 1$  in  $b_1 = b_2 = 0$  in je prikazan na sliki 4.18.

Naš cilj je bila identifikacija GP modela, ki bo opisoval dinamično obnašanje sistema (4.1) na področju, omejenem z vhomom  $u$  med  $u_{\min} = 0$  in  $u_{\max} = 1.5$ . Za vrednotenje dobljenega modela smo primerjali rezultate simulacije GP modela z obnašanjem sistema. Uporabili smo dva različna GP modela:



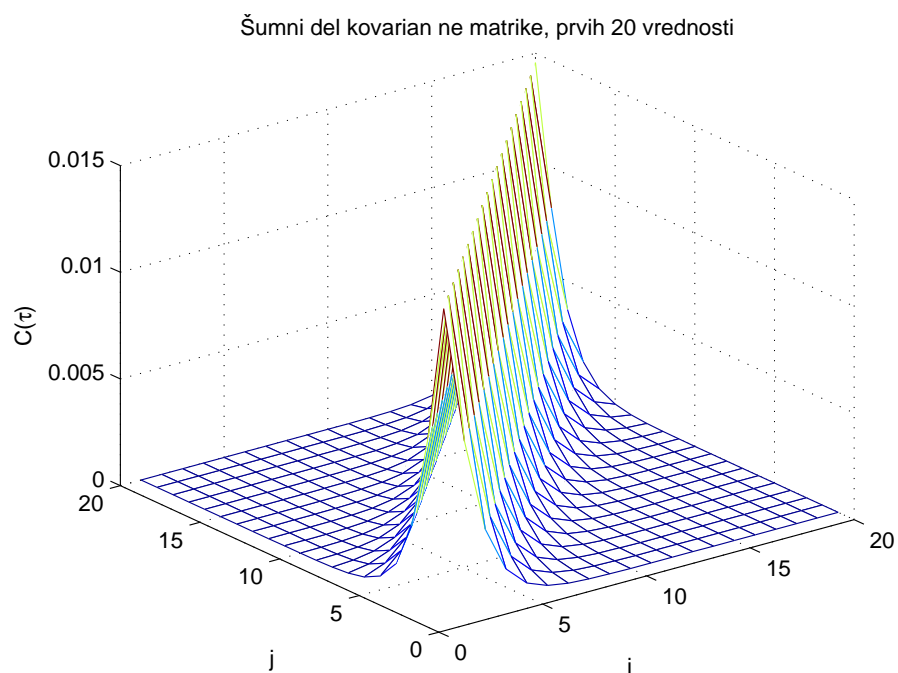
**Slika 4.18:** Model uporabljenega barvnega šuma na izhodu sistema  $e(k)$  v programskem paketu Matlab/Simulink

- GP model brez predznanja o šumu in
- GP model z vključenim predznanjem o modelu šuma na izhodu.

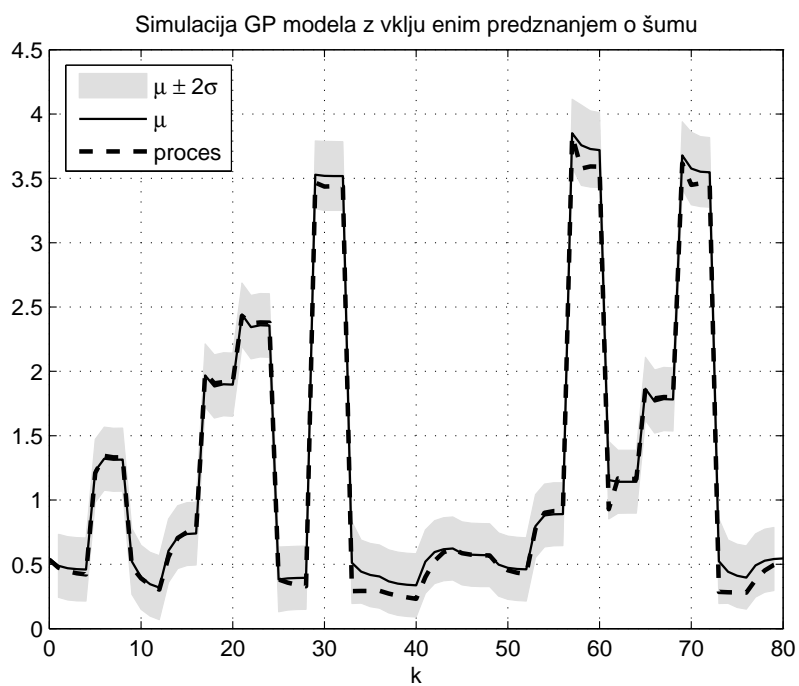
Za učenje GP modelov smo uporabili 202 vzorca, ki smo jih dobili iz dinamičnega odziva modela. Ker nas je zanimalo, kako se odraža vključevanje predznanja o šumu, smo uporabili zaporedne vzorce vhodnega in izhodnega signala, katerih šum je najbolj koreliran. Za vzbujanje sistema je bil uporabljen diskretni psevdonaključni signal (str. 74) s taktom  $T_{sw} = 2$ , katerega amplituda ob spremembi lahko zavzame katero koli vrednost med  $u_{\min}$  in  $u_{\max}$ . Učne izhode GP modela so predstavljali vzorci izhoda sistema  $y(k)$ , odgovarjajoče učne vhode GP modela pa za en korak zakasnjeni vzorci vhoda in izhoda sistema  $[u(k-1), y(k-1)]$ .

Najprej je bil narejen GP model za sistem (4.1) brez vključenega znanja o šumnem procesu. Na izhodu smo predpostavili beli šum in optimirali hiperparametre  $v, w_1, w_2$  ter  $v_0$ . Za učenje drugega modela, tj. GP modela z vključenim znanjem o šumu na izhodu procesa, so bili uporabljeni isti vzorci kot za učenje prvega modela. Predznanje o šumu smo opisali s šumnim delom kovariančne matrike  $\mathbf{K}_n$  (slika 4.19), kot je bilo predhodno opisano. Ker smo šumni proces poznali, hiperparametra  $v_0$  ni bilo.

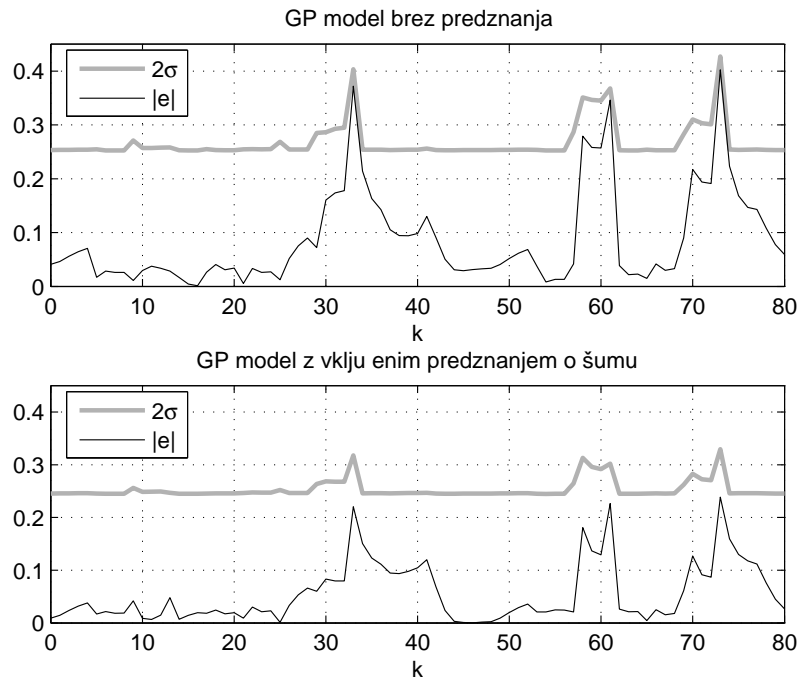
Za ovrednotenje obeh modelov smo modela vzbujali z diskretnim psevdonaključnim signalom s taktom  $T_{sw} = 4$  in primerjali rezultate simulacije z odzivom sistema. Rezultate vrednotenja modela z vključenim predznanjem vidimo na sliki 4.20. Na sliki 4.21 je primerjava med napako GP modelov z in brez vključenega predznanja o modelu šuma. Zaradi ilustrativnosti so prikazani rezultati le za manjši del tega signala. Opazno je boljše obnašanje (manjša napaka pri simulaciji) drugega modela, saj prvi manj ustrezno modelira šum na izhodu. Vrednosti cenilk SE in LD, zbrane v tabeli 4.4, potrjujejo boljše obnašanje drugega modela.



Slika 4.19: Šumni del kovariančne matrike  $\mathbf{K}_n$



Slika 4.20: Vrednotenje GP modela z vključenim predznanjem o šumu na signalu za vrednotenje



Slika 4.21: Primerjava napak simulacij obeh GP modelov

Tabela 4.4: Vrednosti cenilk za rezultata simulacije GP modelov z vključenim predznanjem o šumu in brez njega

Cenilka	“Brez predznanja”	“Vključeno predznanje”
SE	0.0106	0.0045
MRSE	0.061	0.040
LD	-0.88	-1.03

## Poglavje 5

# Praktični primeri uporabe GP modelov

### 5.1 Uvod

V prejšnjih poglavjih smo predstavili GP model in ilustrirali njegovo uporabo na simulacijskih primerih, izjema je bil primer vrednotenja GP modela za napovedovanje rasti alg v podpoglavju 3.8. V tem poglavju bomo prikazali uporabo GP modela na primerih, pri katerih so podatki pridobljeni iz meritev procesov. Ti predstavljeni praktični primeri uporabe GP modela, predvsem za opis dinamičnih sistemov, bodo najboljše prikazali njegove lastnosti in uporabnost.

#### Lastnosti GP modela, pomembne za praktično uporabo

Povzemimo poglobitve (že omenjene) dobre lastnosti GP modela za identifikacijo nelinearnih dinamičnih sistemov:

- Poleg napovedi izhoda sistema, ki ga opisuje, daje GP model na izhodu tudi mero zaupanja v napovedan izhod. Ta je izražena z varianco in je odvisna od učnih podatkov.
- GP model za učenje potrebuje razmeroma majhno število učnih podatkov, njihovo potrebno število je odvisno od kompleksnosti opisovanega sistema.
- V model je mogoče vključevati predznanja različnih vrst, kar je v skladu z napotkom, da znanje, ki ga imamo na voljo, koristno uporabimo.
- Model je enostaven za uporabo. Predvsem je enostavna izbira njegove strukture. To lahko naredi pristop privlačen kot orodje za identifikacijo.

Izmed omejitev GP modela na prvo mesto postavimo neparametričnost modela, ki omejuje njegovo uporabo. Tipičen primer je načrtovanje vodenja, saj je možna uporaba neparametričnega modela samo v omejenem številu načrtovalnih metod. Primer take metode je prediktivno vodenje. Za zdaj sta problematični še:

- počasnost učenja, ki se zna z nadaljnjim razvojem metode pohitriti (nekatero izmed možnih rešitev so opisane v razdelku 3.2.1 na str. 27) in
- razširjenost uporabe GP modela, ki pa se tudi izboljšuje, saj je na voljo vedno več programske opreme<sup>1</sup>, literature in primerov uporabe. Predlog našega koncepta programskega modela za identifikacijo dinamičnih sistemov se nahaja v Dodatku F.

### Možna uporaba GP modela

GP model in njegove izpeljanke lahko uporabimo bolj ali manj za vse naloge, zaradi katerih sisteme sploh identificiramo:

- napovedovanje odziva dinamičnih sistemov, naj bo to enokoračno napovedovanje (biološki sistemi, astronomija, poraba električne energije), večkoračno napovedovanje ali simulacija;
- načrtovanje vodenja;
- zaznavanje napak;
- itd.

uprabimo pa ga lahko npr. tudi za glajenje signalov oz. podatkov, ki vsebujejo šum.

### Primeri, ki bodo uporabljeni za predstavitev delovanja GP modela

V tem poglavju je predstavljena uporaba GP modela na nekaj praktičnih problemih. Vse predstavljene primere smo poskušali izbrati tako, da smo prikazali ali nov način modeliranja dinamičnih sistemov ali da je šlo za prvo uporabo GP modelov v ta namen.

Najprej je predstavljena uporaba GP modela za glajenje identificiranega frekvenčnega odziva, ki predstavlja uporabo GP modela na statičnem

<sup>1</sup>Npr.: The Gaussian Processes web site: <http://www.gaussianprocess.org/>

regresijskem problemu. Sledi primer identifikacije procesa nitrifikacije v biološkem čistilnem procesu, v katerem je korak po koraku predstavljen postopek identifikacije z GP modelom, poudarek pa je namenjen njegovemu vrednotenju. V četrtem podpoglavju je predstavljena identifikacija FSGP modela sistema dveh posod, ki je nato uporabljen za načrtovanje vodenja. V petem podpoglavju je predstavljena identifikacija FSGP modela procesa priprave plina, ki je prav tako uporabljen za načrtovanje vodenja. V šestem podpoglavju je na sistemu dveh posod predstavljen primer zaznavanja napak z uporabo LMGP modela, pri čemer pomembno vlogo igra varianca napovedi modela. Na koncu so predstavljeni nasveti, kdaj uporabiti katero izmed izpeljank GP modela.

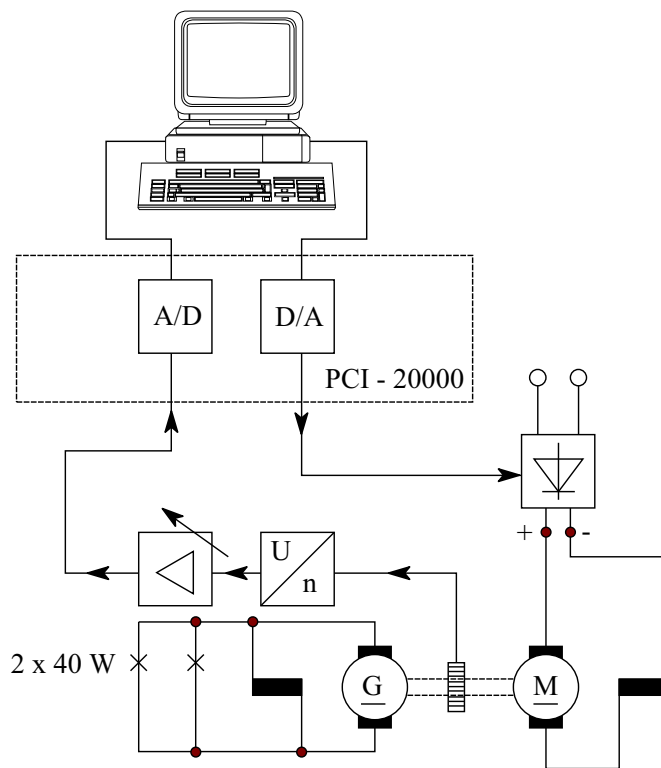
## 5.2 Glajenje odzivov z GP modelom

V tem razdelku bo predstavljena uporaba GP modela za glajenje podatkov, ki spada med statične regresijske probleme. Za njihovo reševanje je uporaba GP modela že dolgo znana [81], na primeru smo jo predstavili v podpoglavju 2.2. Uporabili jo bomo za glajenje frekvenčnega odziva nelinearnega dinamičnega sistema.

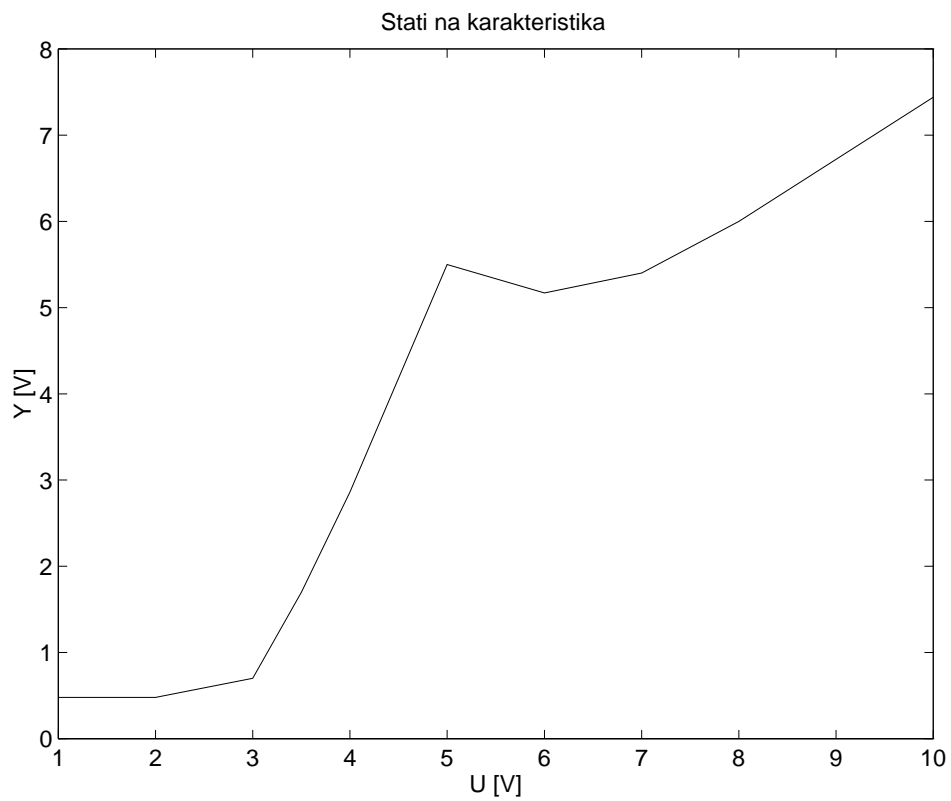
V [43] je prikazana identifikacija električne pilotne naprave motor-generator [21] proizvajalca ELWE. Gre za laboratorijsko napravo, ki je namenjena eksperimentalnemu delu na področju načrtovanje vodenja. Njeno shemo skupaj z opremo za zajemanje podatkov prikazuje slika 5.1.

Enosmerni motor moči 100 W s serijskim vzbujanjem je mehansko povezan s tuje vzbujanim enosmernim generatorjem. Breme generatorja sta dve 40 W žarnici (220 V). Identificiran je bil proces, pri katerem kot vhod nastopa napetost na sponkah motorja in kot izhod število vrtljajev oziroma iz stališča meritev model sistema, pri katerem vhod predstavlja napetost na sponkah motorja  $U$ , izhod pa napetost iz merilnega pretvornika vrtljajev  $Y$ . Tiristorski pretvornik na vhodu sistema omogoča vodenje s procesnim računalnikom. Izhod iz merilnika vrtljajev daje na izhodu 1 V napetosti pri 1000 vrtljajih na minuto, ki jo peljemo še prek merilnega ojačevalnika, tako da sta vhodni in izhodni signal v območju od 0 do 10 V. Tako tiristorski pretvornik kot ojačevalnik sta del spremljevalne opreme sistema motor-generator istega proizvajalca. Za digitalno analogno in analogno digitalno pretvorbo pri zajemanju in dovažanju signalov je bil uporabljen pretvornik PCI-2000 proizvajalca Burr-Brown.

Izmerjeno statično karakteristiko procesa prikazuje slika 5.2. Sistem želimo identificirati v delovni točki, določeni z vhodno napetostjo  $U = 8.5$  V, v kateri je sistem dokaj linearen.

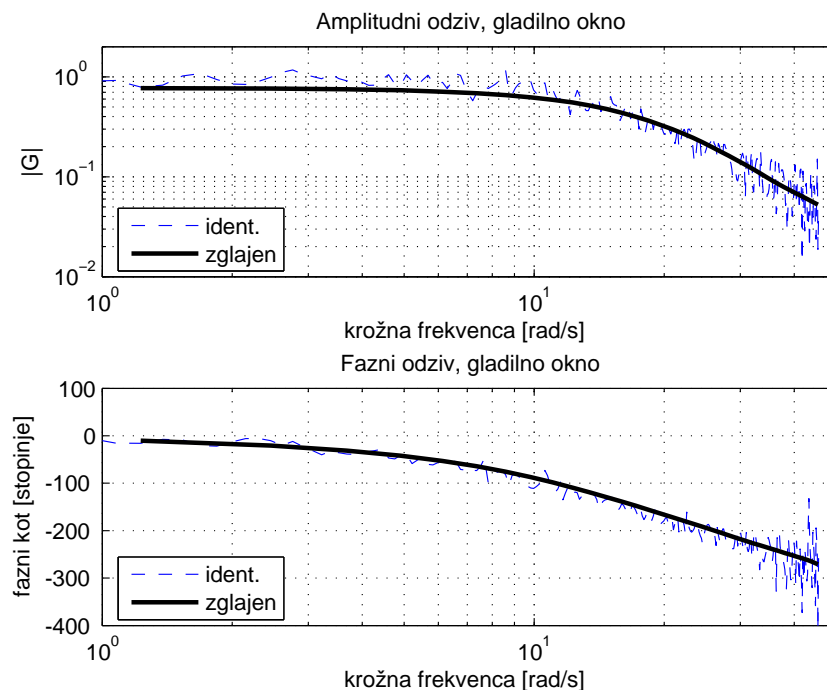


Slika 5.1: Shema sistema motor-generator in opreme za zajemanje podatkov



Slika 5.2: Statična karakteristika sistema, ki ga identificiramo

Ena izmed uporabljenih metod identifikacije v [43] je bila tudi računanje frekvenčnega odziva v izbrani delovni točki s spektralno analizo [69]. Ta metoda je zelo občutljiva na prisotnost šuma, varianca njene ocene pa je pri velikih premikih med signaloma  $\tau$  velika. Oba problema sta rešljiva z glajenjem, za katerega je bila v [43] uporabljena metoda glajenja s Hammingovi okni [69]. Primerjava med izračunanim in s Hammingovi okni zglajenim frekvenčnim odzivom je prikazana na sliki 5.3.



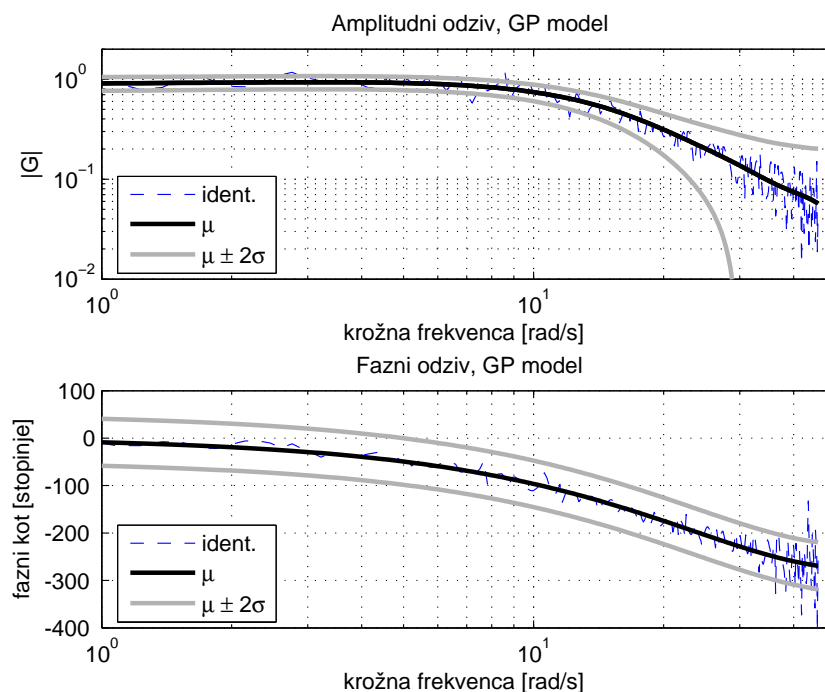
**Slika 5.3:** Amplitudni in fazni odziv brez glajenja (črtkana krivulja) in z glajenjem s Hammingovi okni (polna krivulja) [43]

Rezultati so videti razmeroma v redu, a uporaba glajenja s Hammingovim oknom ni nujno enostavna, saj moramo izbrati primerno širino okna za glajenje [69].

Odločili smo se, da bomo kot alternativni način za glajenje izmerjenih odzivov uporabili dva GP modela – prvega za amplitudni in drugega za fazni odziv. Učni množici za GP modela sta identificirani amplitudni in fazni odziv sistema. Vhodni del učne množice predstavljajo vrednosti krožnih frekvenc, izhodni del pa vrednosti amplitudnega oz. faznega odziva pri ustreznih vhodnih vrednostih.

Ker pričakujemo gladka odziva, za kovariančno funkcijo izberemo Gaussovo kovariančno funkcijo (2.4). Za učenje vsakega od modelov uporabimo 297 točk; vrednosti frekvenc, ki nastopajo kot vhodi, so med 0.15 in 45 Hz. Hiperparametre izberemo z metodo največje podobnosti.

Izhod iz prvega GP modela je zglajena vrednost amplitudnega odziva, medtem ko je izhod iz drugega modela zglajena vrednost faznega odziva, obe pri ustrezni vrednosti frekvence, podane na vходу v GP modela. Za predstavitev rezultatov za testne vrednosti vhoda v GP model izberemo kar iste vrednosti frekvenc, ki so bile uporabljene že za učenje GP modelov. Rezultati glajenja so predstavljeni na sliki 5.4. Iz slike lahko razberemo, da so rezultati glajenja dobri, in če pogledamo npr. amplitudni odziv v spodnjem delu frekvenčnega spektra, so boljši kot rezultati, dobljeni z glajenjem s Hammingovi okni, saj boljše ujamejo srednjo vrednost izmerjenega amplitudnega odziva. Poleg dobrih srednjih vrednosti zglajenega frekvenčnega in faznega odziva dobimo, kot je za GP modele v navadi, še spremljajočo mero negotovosti napovedanih odzivov.



**Slika 5.4:** Amplitudni in fazni odziv brez glajenja (črtkana krivulja) in z glajenjem z dvema GP modeloma (polna krivulja)

Če bi imeli več med seboj različnih amplitudnih in faznih odzivov istega sistema, npr. odzive v različnih delovnih točkah ali odzive časovno spremenljivega sistema, bi lahko na prikazani način z GP modelom prikazali povprečen zglajeni odziv s pasom negotovosti, v katerem bi bila vsebovana odstopanja posameznih odzivov od povprečja.

Prikazani statični nelinearni regresijski problem je bil za ilustracijo glajenja kot enega izmed bolj znanih načinov uporabe GP modela. Postopek glajenja z GP modelom je preprost, potrebna je le izbira kovariančne funkcije, za katero navadno lahko izberemo kar Gaussovo kovariančno funkcijo (2.4), izbira kakšnih dodatnih

parametrov ni potrebna. Dodatna ugodnost pri uporabi GP modela za glajenje je dano zaupanje v napoved, ki da dodatno informacijo o rezultatih. Tako je model koristen tudi, ko so podatki, ki jih želimo zgladiti, razporejeni nehomogeno, tj. ko imamo področja z majhno in veliko gostoto podatkov. V takem primeru nas varianca napovedi opozori na področja, kjer je napoved (verjetno) slabša. GP model ni namenjen ekstrapoliranju, na kar bi nas prav tako opozorila povečana varianca.

### 5.3 Identifikacija biološke čistilne naprave za napovedovanje odziva

GP model je že bil uporabljen za identifikacijo različnih dinamičnih sistemov iz simuliranih podatkov (npr. [45, 47, 58]) in tudi iz izmerjenih podatkov: laboratorijske naprave treh posod [4, 7] in procesa priprave plina [61, 62], ni pa še bil uporabljen na področju identifikacije biotehnoloških sistemov.

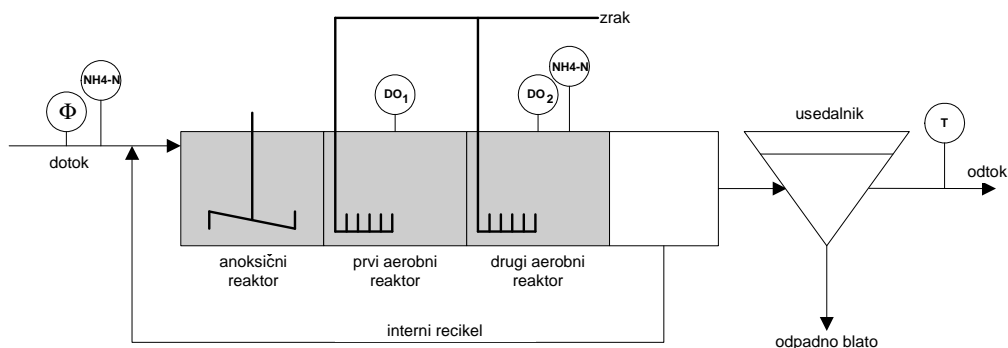
Biotehnološki sistemi so težki za modeliranje, saj ni nujno, da poznamo vse procese, ki sistem sestavljajo, prav tako so procesi dostikrat nelinearni (nasičenja, nelinearne odvisnosti spremenljivk stanj itd.). Dostikrat ne poznamo in/ali ne moremo meriti vseh njegovih vplivnih spremenljivk, ki jih zato tudi ne moremo vključiti v model, poleg tega pa meritve procesnih spremenljivk vsebujejo razmeroma veliko šuma. Vse to otežuje identifikacijo takih procesov, ki zato predstavljajo primerno študijo za uporabo GP modela.

V tem podpoglavju bomo predstavili identifikacijo procesa nitrifikacije v čistilni napravi za odpadne vode z GP modelom [10, 12]. Predstavljen bo postopek identifikacije dinamičnih sistemov z GP modelom, podrobneje razčlenjen in opisan v tretjem poglavju, pri čemer bo poseben poudarek namenjen vrednotenju.

#### Opis procesa

Zanima nas model procesa nitrifikacije v biološki čistilni napravi za odpadne vode Domžale-Kamnik, ki uporablja pritrjeno biomaso (angl. *moving bed biofilm*). Proces je bil v [100] identificiran z različnimi linearnimi in nelinearnimi parametričnimi modeli.

Pilotno napravo sestavljajo anoksični reaktor (176 m<sup>3</sup>), dva aerobna reaktorja (130 m<sup>3</sup> in 117 m<sup>3</sup>), mešalni reaktor (115 m<sup>3</sup>) in usedalnik (600 m<sup>3</sup>) [82, 100]. Shema naprave je prikazana na sliki 5.5.



**Slika 5.5:** Shema procesa nitrifikacije v biološki čistilni napravi za odpadne vode

Glavni proces pri odstranjevanju amonija ( $\text{NH}_4\text{-N}$ ) je nitrifikacija, ki teče v aerobnih reaktorjih. Ta je močno odvisna od koncentracije raztopljenega kisika (angl. *dissolved oxygen*, DO), na katero lahko vplivamo, na proces pa vpliva tudi temperatura. Manj se amonij razkroja v anoksičnem reaktorju, v katerem proces teče ob pomoči prisotnih nitratov. Preostale spremenljivke, ki vplivajo na proces, so še vhodni tok tekočine in koncentracija amonija v njej. Modelirati želimo koncentracijo amonija na izhodu iz procesa.

Namen modela je večkoračno napovedovanje, s čimer bi model lahko uporabili npr. za načrtovanje prediktivnega vodenja na podlagi modela ali za nadzorno vodenje na podlagi modela.

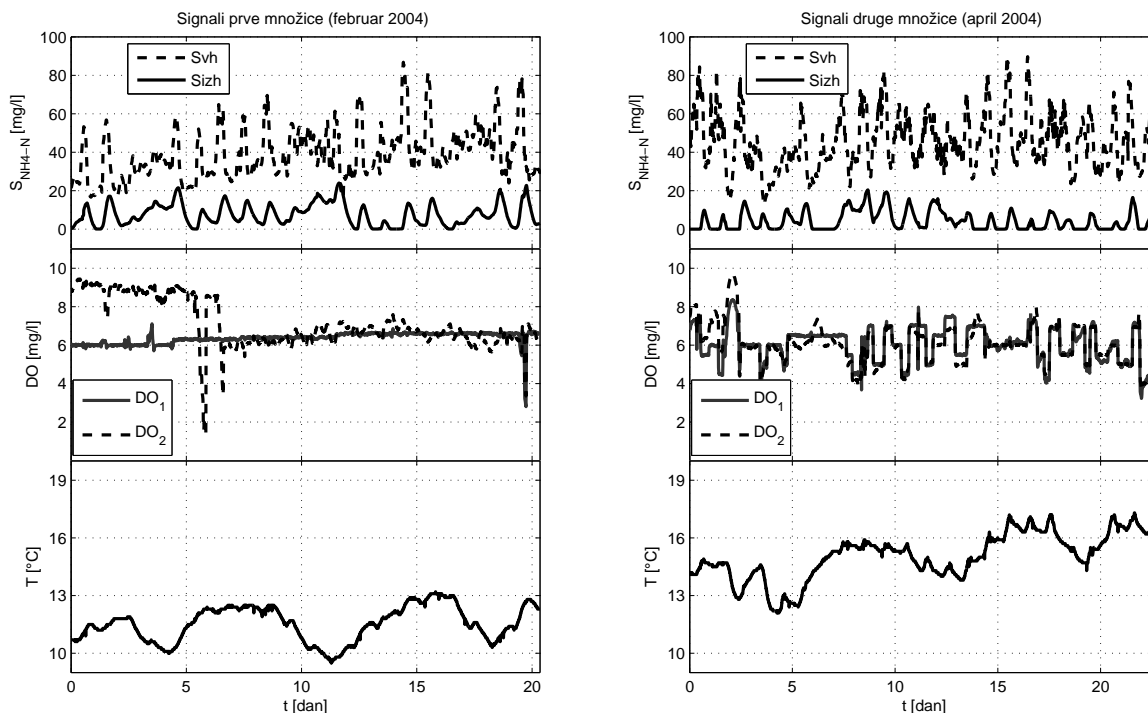
### Podatki za identifikacijo

Na načrtovanje in izvedbo eksperimenta za pridobitev identifikacijskih podatkov nismo imeli vpliva, saj je bil le ta že izveden [100]. Podatki, ki opisujejo obnašanje procesa v čistilni napravi, so bili dobljeni z meritvami v februarju (prva množica) in aprilu 2004 (druga množica). Merjene spremenljivke so zbrane v tabeli 5.1.

**Tabela 5.1:** Merjene procesne spremenljivke za identifikacijo biološke čistilne naprave

Spremenljivka	Oznaka	Enote
vhodna koncentracija amonija	$S_{vh}$	[mg/l]
izhodna koncentracija amonija	$S_{izh}$	[mg/l]
koncentracija kisika v prvem aerobnem reaktorju	$DO_1$	[mg/l]
koncentracija kisika v drugem aerobnem reaktorju	$DO_2$	[mg/l]
temperatura	$T$	$^{\circ}\text{C}$

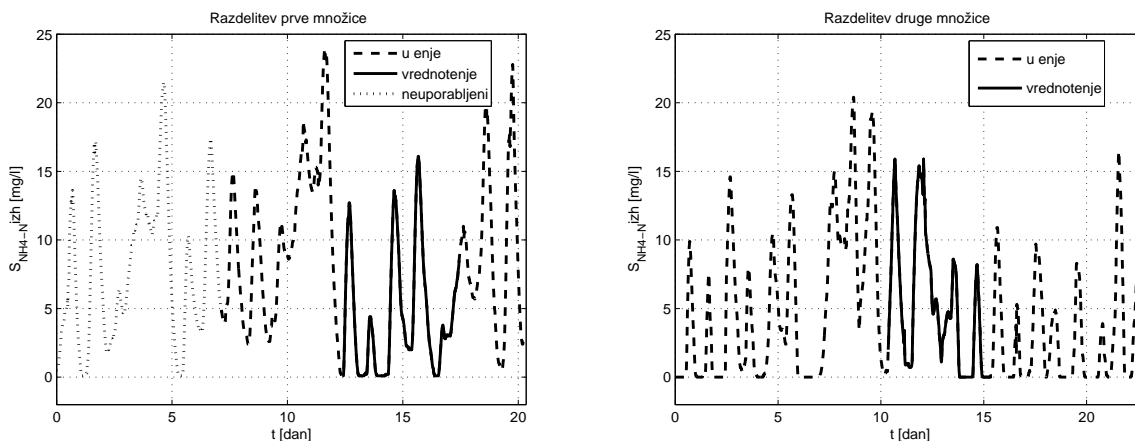
Signali so bili vzorčeni s časom  $T_s = 15$  min. Ker je bil pri vseh meritvah vhodni tok skorajda konstanten, smo njegov vpliv na izhod iz procesa zanemarili. Izmerjeni signali so predstavljeni na sliki 5.6, levo so podatki za februar (prva množica) in desno podatki za april 2004 (druga množica). Zaradi napake na senzorju za merjenje koncentracije raztopljenega kisika v drugem reaktorju  $DO_2$  je ta meritev za prvi teden v prvi množici napačna in ni bila uporabljena.



**Slika 5.6:** Podatki za identifikacijo nitrifikacijskega procesa v čistilni napravi, signali prve množice (levo) in signali druge množice (desno)

Podatke obeh množic meritev smo razdelili na podatke za učenje in na podatke za vrednotenje GP modela. Za učenje so bili uporabljeni vzorci iz začetka in konca obeh množic signalov, za vrednotenje pa petdnevni intervali v vmesnem področju, kot je prikazano na sliki 5.7. S tako razdelitvijo so bili tako učni podatki kot podatki za vrednotenje sestavljeni iz podatkov obeh množic meritev. Vzrok leži v naravi GP modela, ta za dober opis nekega področja potrebuje podatke, ki to področje opisujejo.

Pred učenjem GP modela so bili signali normirani med vrednosti  $-1$  in  $1$ . Vse skupaj smo za učenje GP modela uporabili približno 1200 vzhodno/izhodnih vzorcev.



**Slika 5.7:** Razdelitev podatkov obeh množic na podatke za učenje in podatke za vrednotenje za prvo (levo) in drugo (desno) množico. Prikazan je samo signal, ki opisuje identifikacijsko spremenljivko.

### Postavitev in učenje GP modela

Zaradi že opravljenih meritev in tako že danih identifikacijskih podatkov, sta bili fazi postavitve GP modela in učenja GP modela v identifikacijskem postopku (slika 3.4) združeni.

Zaradi predpostavke, da je opisovani proces nelinearen in stacionaren z gladkim izhodom, izberemo Gaussovo kovariančno funkcijo (2.4). Regresorje izberemo z metodo izbire regresorjev na podlagi vrednotenja (predstavljeno v razdelku 3.4.2 in ilustrirano v razdelku 3.7). Zaradi velikega števila potencialnih regresorjev in relativno majhnega števila podatkov je potrebno metodo modificirati. Pravilno bi bilo v začetnem GP modelu zbrati *use* potencialno uporabne regresorje naenkrat, naučiti model in nato z uporabo ARD lastnosti hiperparametrov iz modela odstraniti neuporabne regresorje. Žal ima modelirani proces veliko število vplivnih vhodov in je število potencialnih regresorjev preveliko za obstoječe število učnih podatkov<sup>2</sup>, zato razbijemo postopek iskanja regresorjev na podlagi vrednotenja na manjše korake. GP modeli s posameznimi skupinami regresorjev so naučeni in ovrednoteni. Boljše GP modele obdržimo. Iz njih z uporabo ARD lastnosti hiperparametrov odstranimo neuporabne regresorje in jim dodamo nove, potencialno uporabne. Tak postopek je daljši in zahtevnejši, saj moramo v principu preizkusiti mnogo kombinacij regresorjev iz celotne množice.

<sup>2</sup>Tudi če bi imeli na voljo dovolj učnih podatkov, bi učenje tako obširnega modela zahtevalo znatne količine računalniškega časa zaradi potrebe po invertiranju velike kovariančne matrike, enačbi (2.10) in (2.11), v vsakem koraku optimizacije [91]. Tako je bil pri približno 1200 učnih točkah in desetih regresorjih na računalniku P4 1.8 GHz z 1 GB spomina čas trajanja optimizacije 175 min za 1000 tekov, čas vrednotenja s simulacijo pa 98 min.

Na podlagi predznanja iz [100] in njegovih avtorjev smo za začetni GP model izbrali regresorje:  $[S_{\text{izh}}^{1,2,3}, S_{\text{vh}}^4, T^2, DO_1^3, DO_2^1]$ , kjer z  $\xi^i$  označimo spremenljivko  $\xi$ , zakasnjeno za  $i$  vzorcev glede na želeni izhod  $S_{\text{izh}}(t)$ :  $\xi^i = \xi(t - iT_s)$ . Za vrednotenje in primerjavo različnih identificiranih modelov so bili uporabljeni vizualni pregled rezultatov in ovrednotenje z istimi cenilkami kot pri identifikaciji modela bioreaktorja v podpoglavju 3.7, tj. negativni logaritem verjetnosti učnih podatkov LL (3.23), povprečna relativna kvadratna napaka MRSE (3.21) in logaritem napovedane gostote napake LD (3.22). V tabeli 5.2 si lahko ogledamo vrednosti cenilk za nekatere izmed identificiranih GP modelov za ilustracijo smeri iskanja strukture modela. Iz posameznih modelov so bili z uporabo ARD lastnosti hiperparametrov odstranjeni regresorji z majhnim prispevkom informacije.

**Tabela 5.2:** Identifikacija in vrednotenje GP modelov z različnimi skupinami regresorjev – vrednosti cenilk

Regresorji					Ident. podatki LL	Podatki za vrednotenje*	
$S_{\text{izh}}$	$S_{\text{vh}}$	$T$	$DO_1$	$DO_2$		MRSE	LD
3,2,1	4	2	2	1	-3590	1.950	148
3,2,1	4	4	2,4	1,5	-3617	0.828	271
3,2,1	3,4	2,4	2,4	1,2	-3643	0.325	114
3,2,1	3,4,7	2,4	2,4	1,2,4	-3649	0.321	114
†3,2,1	3,4	2,6	2,3	2,4	-3649	0.313	108
♠ <b>3,1</b>	<b>3</b>	<b>6</b>	<b>2,3</b>	<b>2,4</b>	<b>-3640</b>	<b>0.295</b>	<b>96</b>
3,2,1	3,4,7	2,4,6	2,3,4	2,4	-3650	0.316	110
3,2,1	3,4,7	2,4,6	2,4,6	1,2,4	-3650	0.319	112
3,2,1	3,4,9	2,6,9	2,3	1,2,4	-3657	0.318	113

Opombe:

\* Združeni podatki za vrednotenje iz obeh množic.

† Izbran GP model.

♠ Izbran GP model z odstranjenimi neinformativnimi regresorji.

S tem postopkom smo dobili zadovoljiv model, v tabeli 5.2 označen s †, z regresorji:  $[S_{\text{izh}}^{1,2,3}, S_{\text{vh}}^{3,4}, T^{2,6}, DO_1^{2,3}, DO_2^{2,4}]$ . Vrednosti njegovih hiperparametrov so prikazani v tabeli 5.3. Lahko opazimo, da so trije hiperparametri glede na preostale razmeroma majhni, tj. hiperparametri, ki ustrezajo regresorjem  $S_{\text{izh}}^2$ ,  $S_{\text{vh}}^4$  in  $T^2$ , manjši od  $10^{-4}$ . Ustrezni regresorji so bili odstranjeni iz modela, s čimer smo dobili končni model, v tabelah 5.2 in 5.3 označen z ♠, z regresorji:  $[S_{\text{izh}}^{1,3}, S_{\text{vh}}^3, T^6, DO_1^{2,3}, DO_2^{2,4}]$ . Zaradi odstranitve manj pomembnih regresorjev se je vrednost cenilke LL za okrnjeni model sicer povečala, vendar je model glede na vrednosti cenilk MRSE in LD najboljši izmed identificiranih.

**Tabela 5.3:** Vrednosti hiperparametrov za izbran GP model

Zakasnitev	$S_{\text{izh}}$			$S_{\text{vh}}$		$T$		$DO_1$		$DO_2$		$v$	$v_0$
	1	2	3	3	4	2	6	2	3	2	4		
Model 1 <sup>†</sup>	0.089	$1.02 \cdot 10^{-6}$	$1.45 \cdot 10^{-2}$	$3.0 \cdot 10^{-3}$	$8.3 \cdot 10^{-5}$	$4.4 \cdot 10^{-5}$	$2.1 \cdot 10^{-4}$	$4.1 \cdot 10^{-4}$	$1.10 \cdot 10^{-3}$	$1.33 \cdot 10^{-3}$	$3.3 \cdot 10^{-4}$	6.2	$1.33 \cdot 10^{-4}$
Model 2 <sup>♣</sup>	0.2	×	$2.7 \cdot 10^{-2}$	$2.9 \cdot 10^{-3}$	×	×	$1.07 \cdot 10^{-2}$	$6.8 \cdot 10^{-4}$	$1.27 \cdot 10^{-3}$	$3.4 \cdot 10^{-3}$	$8.8 \cdot 10^{-4}$	2.6	$1.32 \cdot 10^{-4}$

Opombe:

Vrednosti hiperparametrov za normalizirane podatke:

<sup>†</sup> Izbran GP model.

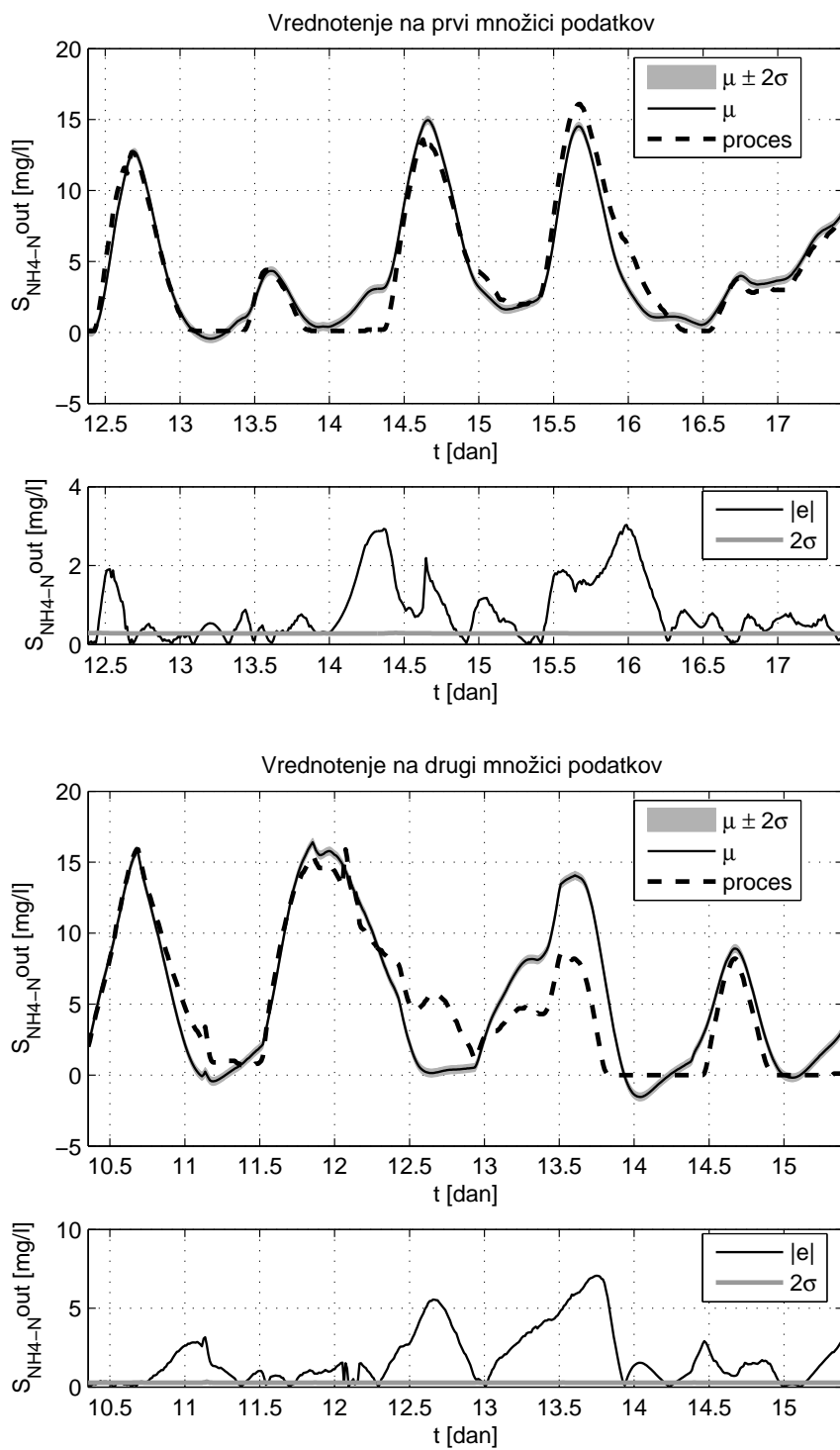
<sup>♣</sup> Izbran GP model z odstranjenimi neinformativnimi regresorji.

## Vrednotenje

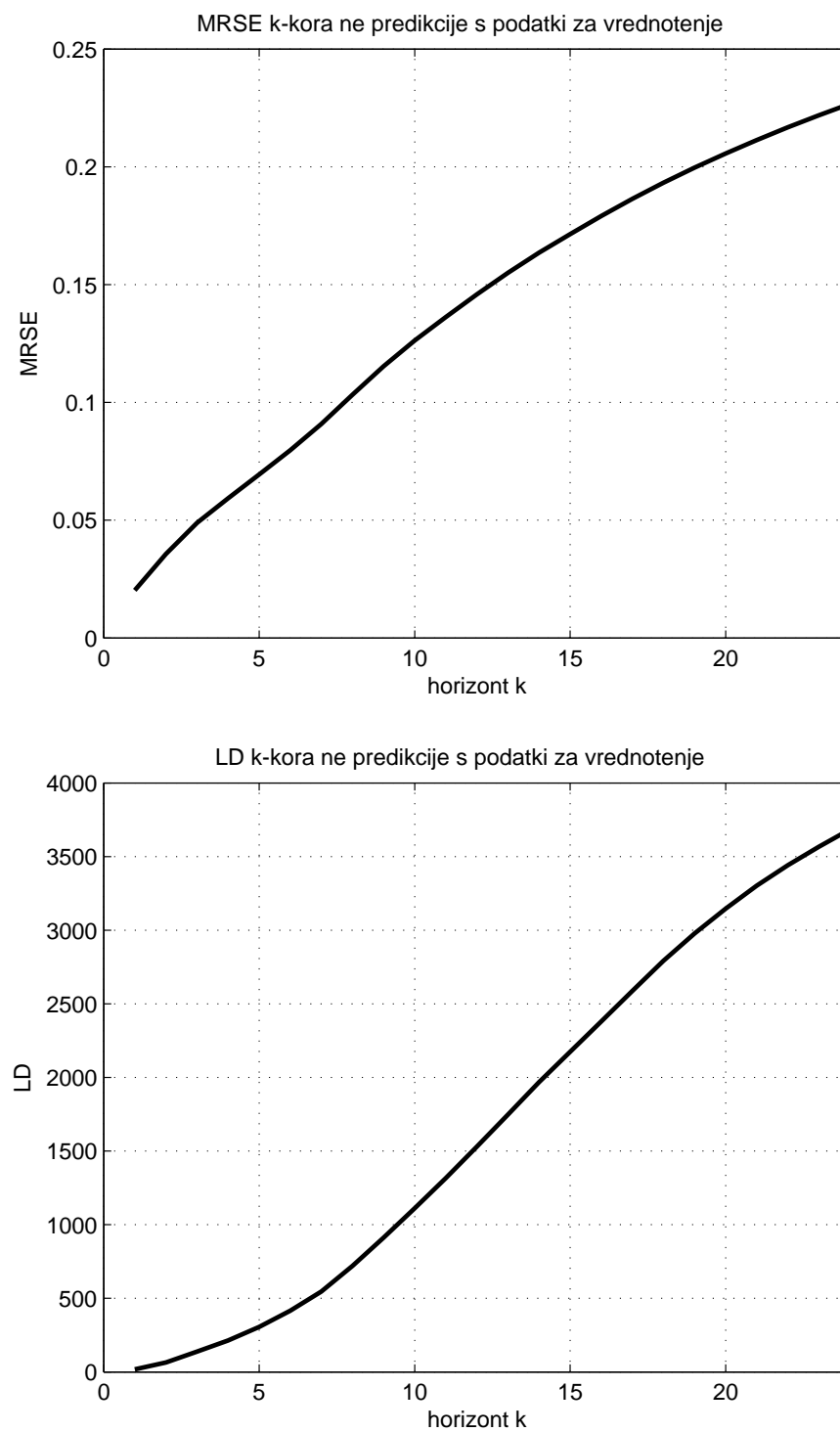
Za kvalitativno vrednotenje je bila uporabljena “naivna” simulacija na obeh množicah podatkov za vrednotenje. Rezultati so prikazani na sliki 5.8. Vrhovi in doline odziva procesa in modela se ujemajo, iz česar lahko sklepamo na zadovoljivo dinamično obnašanje. Relativno nizka ocena variance šuma procesa  $v_0 = 0.0191$  in ozka pasova 95% zaupanja glede na napako napovedovanja nakazujejo, glede na prikazano v podpoglavju 3.7, da nekatere odvisnosti procesa od zunanjih vplivov niso zajete.

Primerjava rezultatov simulacije GP modela z rezultati iz [100], kjer je imel model procesu primerljivo dinamično obnašanje in MRSE okoli 0.45, kaže na rahlo prednost GP modela. Tu je potrebno omeniti, da je razdelitev podatkov na podatke za učenje in podatke za vrednotenje drugačna kot v [100], v katerem so bili za učenje modela uporabljeni celotni podatki iz druge množice, za vrednotenje pa celotni podatki iz prve množice. Za modeliranje, pri katerem kombiniramo teoretično modeliranje z identifikacijo tako, da struktura modela odseva predpostavljene fizikalne, kemične in druge odnose med spremenljivkami, pomembnost izbire podatkov za učenje ne pride tako do veljave, saj struktura modela odseva predpostavljene odnose med regresorji na celotnem področju delovanja. V GP modelu pa razmerja med podatki niso dana eksplicitno in model mora na področjih, neopisanih s podatki, ekstrapolirati.

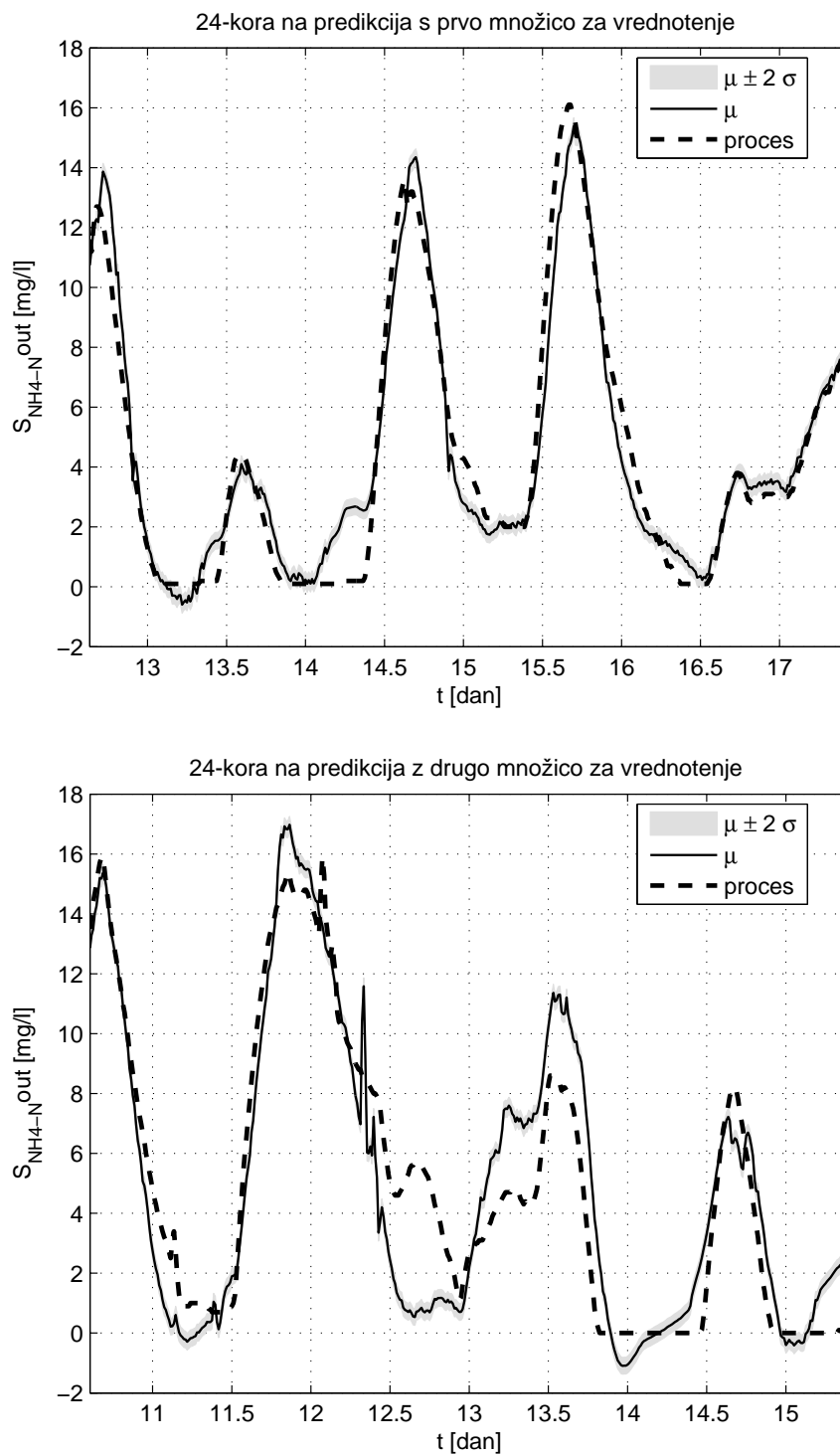
Parametrični model iz [100] je bil identificiran za uporabo pri načrtovanju prediktivnega vodenja. Za ovrednotenje namembnosti GP modela je bil model uporabljen tudi za večkoračno napovedovanje. Odvisnosti cenilk MRSE in LD od horizonta  $h = 1$  naprej sta prikazani na sliki 5.9. Rezultati 24-koračne predikcije na obeh množicah za vrednotenje so prikazani na sliki 5.10. Rezultati 24-koračne predikcije, tako kot rezultati simulacije, z ozkim pasom zaupanja in veliko vrednostjo cenilke LD sugerirajo obstoj nemodeliranih odvisnosti procesa. Kljub vsemu je srednja vrednost napovedi, izražena skozi vrednost cenilke MRSE, sprejemljiva. Ker je GP model v luči kvantitativnih rezultatov rahlo boljši od modela iz [100], lahko ocenimo, da bi bil tudi GP model uporaben za vodenje, tako kot je bilo to ocenjeno in prikazano v [100].



**Slika 5.8:** Vrednotenje s simulacijo na podatkih iz prve množice (zgoraj) in druge množice (spodaj)



**Slika 5.9:** Vrednost cenilke MRSE (zgoraj) in LD (spodaj) večkoračne predikcije v odvisnosti od horizonta  $k$



**Slika 5.10:** Večkoračna predikcija pri horizontu  $k = 24$  ( $t = 6$  h) za prvo (zgoraj) in drugo množico podatkov za vrednotenje (spodaj)

\* \* \*

Na predstavljenem primeru je bila prikazana metodologija identifikacije z GP modelom s poudarkom na vrednotenju. Dobljen GP model je uporaben kljub omejeni informaciji, ki je na voljo za njegovo učenje. Pomanjkanje informacije in prisotnost šuma v meritvah sta pogosta pojava pri modeliranju bioloških in biotehnoloških procesov, zato je lahko predstavljeni primer tudi spodbuda za pogostejšo uporabo GP modela kot orodja za identifikacijo tovrstnih procesov.

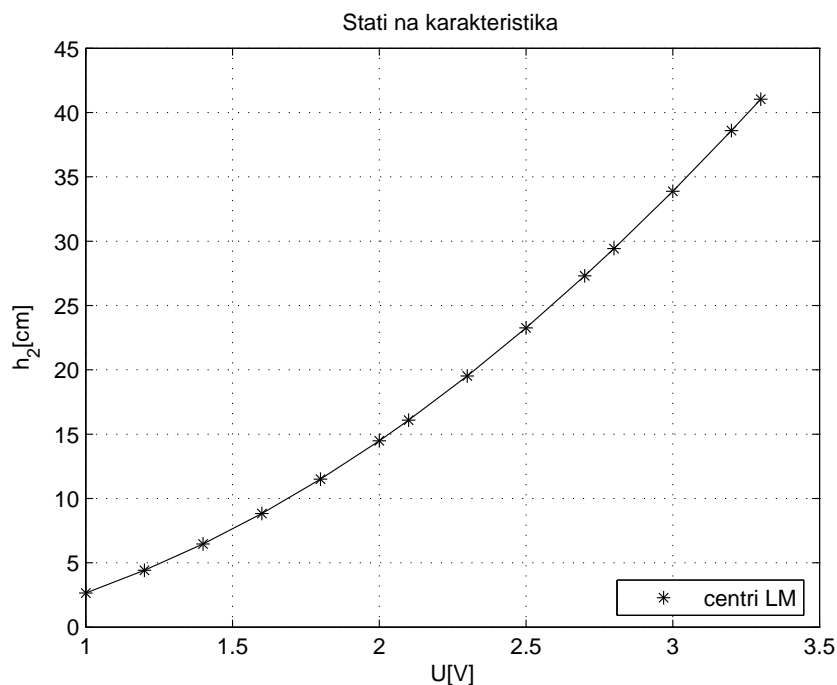
## 5.4 Identifikacija sistema dveh posod na hidravlični modelni napravi z uporabo FSGP modela za načrtovanje vodenja

V tem poglavju bomo identificirali model sistema dveh posod hidravlične modelne naprave [19, 89], predstavljene v Dodatku E.1, na podlagi katerega bomo načrtovali preprosto vodenje s kompenzacijskim regulatorjem, ki smo ga uporabili že v [9]. Ker za načrtovanje takega regulatorja potrebujemo parametrični model, se odločimo za modeliranje procesa s FSGP modelom. Seveda obstajajo tudi načini za načrtovanje vodenja na podlagi nelinearnih modelov, kamor spada tudi običajni GP model, najdemo jih npr. v [1, 2], vendar imamo z uporabo parametričnih modelov raznovrstnejše možnosti načrtovanja. Dodajmo še, da poudarek tega primera ni na metodi načrtovanja vodenja, ampak na metodi modeliranja procesa.

V nadaljevanju si bomo ogledali identifikacijo sistema s FSGP modelom in njegovo vrednotenje, nato pa bomo dobljeni model uporabili za načrtovanje vodenja modelne naprave. Načrtovali bomo zaprtizančni sistem, ki bo deloval v glavnem v ravnotežnih točkah sistema, saj je delovanje v neravnotežnem področju zelo omejeno z izvršnimi sistemi modelne naprave.

### Identifikacija sistema dveh posod s FSGP modelom

Sistem dveh posod je nelinearni dinamični sistem, ki ga lahko opišemo z nelinearnim modelom drugega reda (Dodatek E.1). Vhod v sistem predstavlja napetost  $U$  [V], ki krmili hitrost  $\omega_1$  črpalke  $P1$ . Izhod iz naprave je višina stolpca tekočine  $h_2$  [cm], merjena s kapacitivnim senzorjem in pretvorjena iz [V] v [cm]. Za poenostavitev bomo vrednosti navajali brez merskih enot. Na sliki 5.11 je prikazana statična karakteristika sistema. Dinamični sistem pa nima nelinearne le statične karakteristike, ampak je nelinearen tudi dinamično, kar vidimo iz



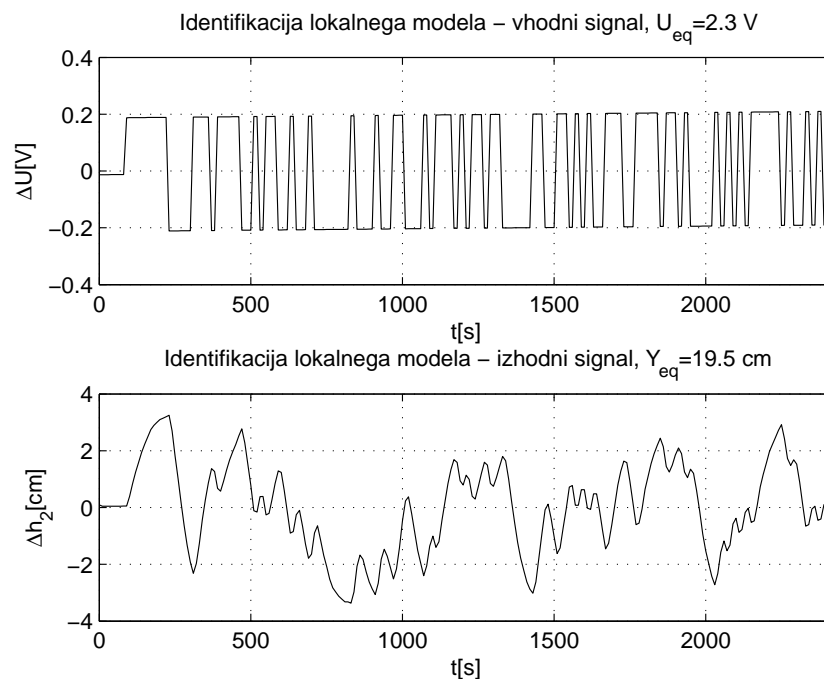
Slika 5.11: Statična karakteristika sistema

enačb (E.1). Področje delovanja je omejeno z višinama tekočine v posodah R1 in R2, ki ne smeta preseči višine obeh posod. S preizkušanjem smo največjo vrednost napetosti na črpalki P1 omejili na  $U_{\max} = 4$ .

Identifikacija dinamičnih sistemov s FSGP modelom je predstavljena v podpoglavju 4.3. Glavna ideja modela je z GP modeli modelirati spremenljive parametre modela z vnaprej določeno, v našem primeru linearno, strukturo. Učni podatki za te GP modele so parametri lokalnih modelov v odvisnosti od delovnega področja, ki je določeno z vrednostmi razvrstilnih spremenljivk. Bogatejši (popolnejši) vektor razvrstilnih spremenljivk izberemo, boljše lahko sklepamo na odvisnost parametrov lokalnih modelov od delovnega področja, toda hkrati potrebujemo precej več podatkov (prekletstvo dimenzije).

Na posameznem (lokalnem) področju bomo ta sistem opisali z lokalnim modelom drugega reda oblike:

$$y(k) = -a_1 y(k-1) - a_2 y(k-2) + b_1 u(k-1) + b_2 u(k-2). \quad (5.1)$$

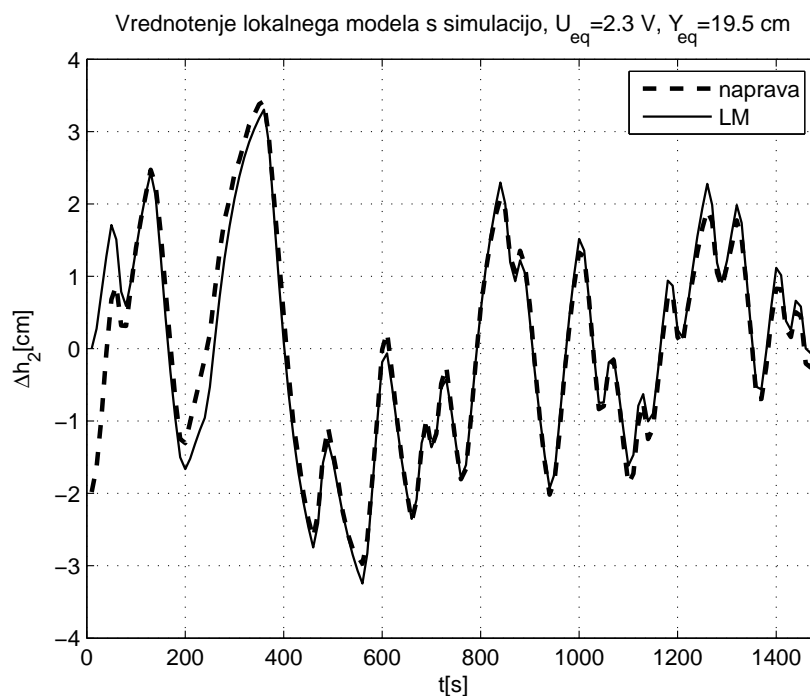


**Slika 5.12:** Signal za identifikacijo lokalnega modela v delovni točki  $(U_{eq}, Y_{eq}) = (2.3, 19.4)$

#### *Identifikacija parametrov lokalnih modelov na ravnotežni krivulji*

Za določanje vrednosti parametrov lokalnega modela v neki ravnotežni točki  $i$  smo sistem s konstantnim vhodnim signalom  $U_{eqi}$  najprej pripeljali v to točko in ga nato tam vzbuiali s frekvenčno dovolj bogatim signalom, da je bilo mogoče iz meritev izluščiti dovolj informacije za identifikacijo lokalnega modela. Z iterativnim postopkom identifikacije pri različnih časih vzorčenja smo izbrali čas vzorčenja sistema  $T_s = 10$  s. Ko smo sistem pripeljali v delovno točko, smo vrednosti vhodnega signala  $U = U_{eqi}$  dodali psevdo-naključni binarni signal [69] s taktom signala<sup>3</sup>  $T_{sw} = 2T_s = 20$  s in amplitudo  $\Delta U = 0.2$ . Ta amplituda določa področje veljavnosti modela, in ker nas bo zanimalo obnašanje blizu ravnotežne krivulje, smo jo določili ustrezno majhno, tako da je celotno ravnotežno območje pokrito z lokalnimi modeli (slika 5.11 in tabela 5.4). Linearni model je bil v ravnotežni točki identificiran z metodo pomožnih spremenljivk v programskem paketu Matlab (iv4). Na sliki 5.12 sta prikazana vhodni in izhodni signal za točko  $(U_{eq} = 2.3, Y_{eq} = 19.4)$ . Podobno kot identifikacijski signali so bili dobljeni tudi signali za vrednotenje identificiranih linearnih modelov. Na sliki 5.13 je za isto delovno točko prikazano vrednotenje linearnega modela s simulacijo.

<sup>3</sup>Takt signala je najmanjši možni čas, po katerem se amplituda signala spremeni.



**Slika 5.13:** Primerjava odziva lokalnega modela in naprave v delovni točki  $(U_{eq}, Y_{eq}) = (2.3, 19.4)$

S tem postopkom smo identificirali 14 lokalnih modelov. Vrednosti ravnotežnih točk, ki ustrezajo njihovim središčem, njihovi parametri in ustrezne variance parametrov so zbrani v tabeli 5.4. Vidimo, da se variance parametrov posameznih lokalnih modelov med seboj ne razlikujejo veliko.

#### *Tvorjenje in vrednotenje FSGP modela*

Običajno potrebujemo lokalne modele tako v ravnotežnih kot v neravnotežnih področjih. Prve dobimo razmeroma enostavno, pridobivanje lokalnih modelov v neravnotežnih področjih pa je znatno težje. Ker v našem primeru sistem identificiramo z namenom načrtovati vodenje, ki bo zagotavljalo delovanje zaprtozančnega sistema pretežno v ravnotežnem področju delovanja, smo se odločili, da poskusimo model sistema določiti najprej z zmanjšanim vektorjem razvrstilnih spremenljivk, v katerem nastopa samo vrednost izhoda v prejšnjem koraku vzorčenja  $\rho(t - T) = [y(t - T)]$ . Lahko bi se odločili tudi za katerega od drugih regresorjev:  $y(t - 2T)$ ,  $u(t - T)$  ali  $u(t - 2T)$ , vendar se izbira samo regresorja  $y(t - T)$  ponuja iz dveh razlogov:

- dinamične lastnosti sistema so najbolj odvisne od višine tekočine v posodah (enačbe sistema (E.1)) in

- zanimale nas bodo počasnejše spremembe stanj sistema, pri čemer se sistem ne oddalji veliko od ravnotežne krivulje; to dosežemo s počasnim spreminjanjem vhodnega signala. Tako sistem lažje opišemo samo z modeli, ležečimi na ravnotežni krivulji, kar bistveno olajša postopek identifikacije lokalnih modelov.

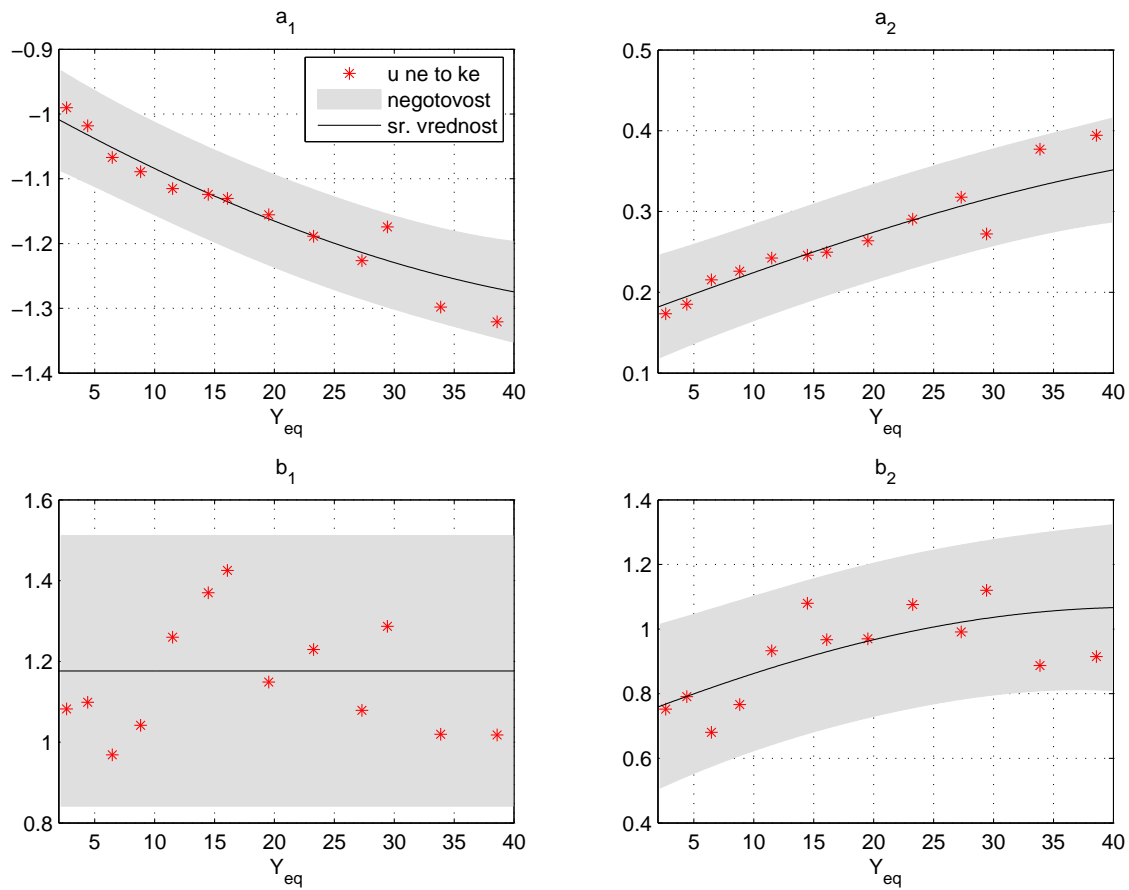
Učenje FSGP modela je učenje v našem primeru štirih GP modelov, ki opisujejo štiri parametre lokalnih modelov  $a_1, a_2, b_1$  in  $b_2$  v odvisnosti od regresorja  $y(t - T)$ . Za kovariančno funkcijo izberemo kar Gaussovo kovariančno funkcijo (2.4), saj predpostavljamo gladko spreminjanje parametrov lokalnih modelov. Učne točke so sestavljene iz parametrov identificiranih lokalnih modelov in ustreznih vrednosti regresorja  $y(t - T)$ . Varianc identificiranih parametrov v učne podatke nismo vključevali, ampak smo prepustili algoritmu za učenje GP modelov, da poišče skupno vrednost variance za podatke.

Vrednosti posameznih hiperparametrov GP modelov po učenju so zbrani v tabeli 5.5. Napovedovanje parametrov skupaj z učnimi podatki si lahko ogledamo na sliki 5.14.

Štiri naučene GP modele smo združili v FSGP model, kot je bilo predstavljeno v podpoglavju 4.3. Algoritem uporabljene funkcije za simulacijo FSGP modela je predstavljen v Dodatku C.3.

**Tabela 5.4:** Parametri identificiranih lokalnih modelov sistema dveh posod, skupaj z ustreznimi variancami

$U_{eq}$	$Y_{eq}$	$a_1$	$\text{var}(a_1)$	$a_2$	$\text{var}(a_2)$	$b_1$	$\text{var}(b_1)$	$b_2$	$\text{var}(b_2)$
1.0	2.7	-0.99	0.0005	0.17	0.0004	1.08	0.0004	0.75	0.0018
1.2	4.4	-1.02	0.0003	0.19	0.0003	1.10	0.0004	0.79	0.0013
1.4	6.5	-1.07	0.0002	0.22	0.0002	0.97	0.0002	0.68	0.0008
1.6	8.8	-1.09	0.0002	0.23	0.0002	1.04	0.0003	0.77	0.0010
1.8	11.5	-1.11	0.0004	0.24	0.0003	1.26	0.0007	0.93	0.0020
2.0	14.5	-1.12	0.0003	0.25	0.0003	1.37	0.0008	1.08	0.0024
2.1	16.1	-1.13	0.0003	0.25	0.0003	1.43	0.0008	0.97	0.0027
2.3	19.5	-1.16	0.0005	0.26	0.0004	1.15	0.0012	0.97	0.0035
2.5	23.3	-1.19	0.0006	0.29	0.0006	1.23	0.0020	1.08	0.0054
2.7	27.3	-1.23	0.0003	0.32	0.0003	1.08	0.0009	0.99	0.0022
2.8	29.4	-1.17	0.0002	0.27	0.0002	1.29	0.0007	1.12	0.0021
3.0	33.9	-1.30	0.0007	0.38	0.0007	1.02	0.0023	0.89	0.0054
3.2	38.6	-1.32	0.0006	0.39	0.0005	1.02	0.0019	0.91	0.0043
3.3	41.0	-1.22	0.0008	0.31	0.0007	1.46	0.0034	1.23	0.0090



Slika 5.14: Napovedovanje parametrov modela z GP modeli skupaj z učnimi podatki

Ker bomo za vhod v modelno napravo uporabljali počasne spremembe vhoda, tak vhodni signal uporabimo tudi za vrednotenje FSGP modela s simulacijo. Primerjava izhoda modelne naprave in odziva FSGP modela na signalu za vrednotenje ter spremljajoče vrednosti spremenljivih parametrov FSGP modela so prikazani na sliki 5.15. Iz rezultatov simulacije lahko sklepamo, da je model za naše potrebe ustrezen. Dinamiko dobro ujame in tudi, če bi prišlo do pogoška v ustaljenem stanju, bi to napako odpravil regulator.

## Vodenje sistema dveh posod z uporabo FSGP modela

### *Načrtovanje vodenja z uporabo FSGP modela*

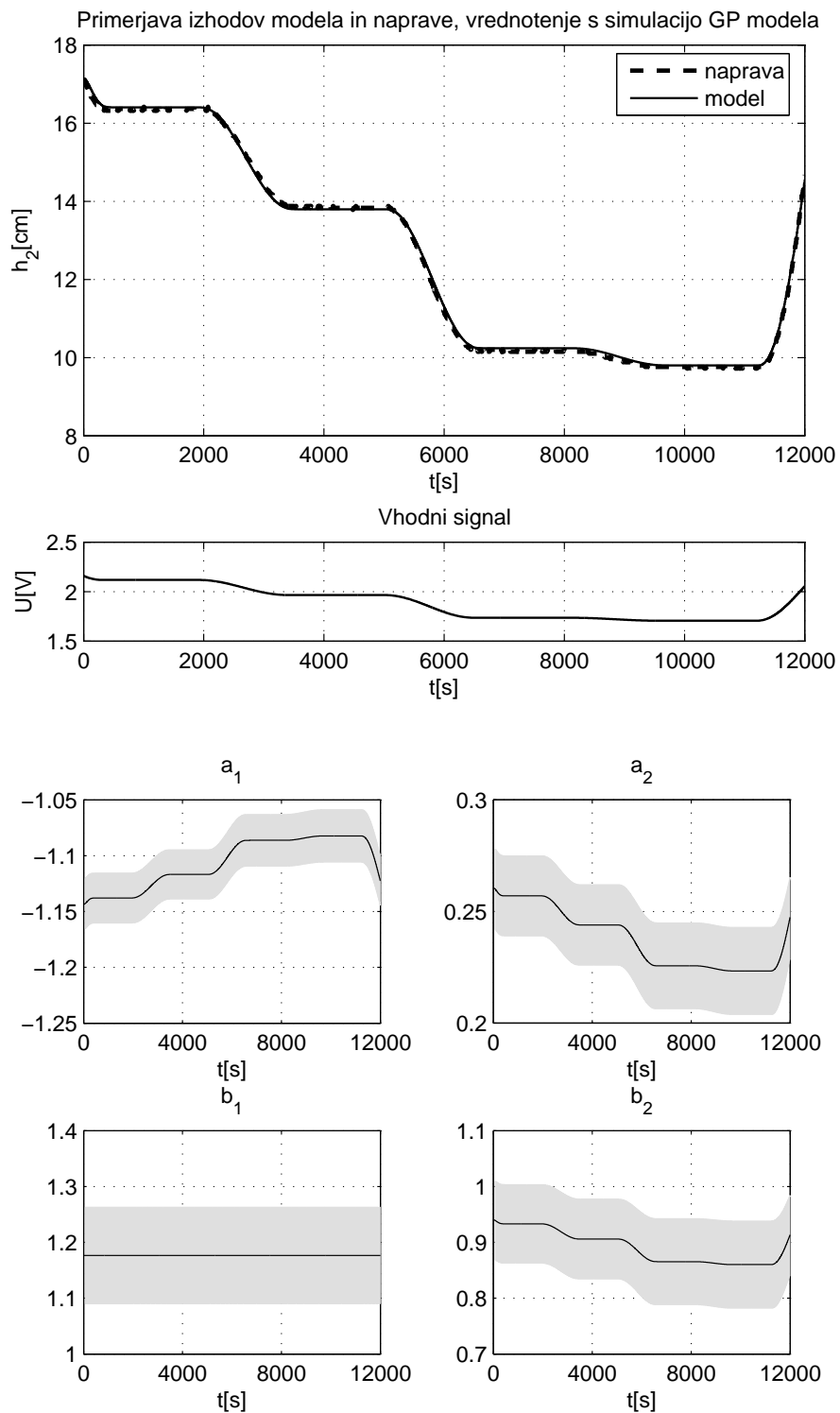
Na podlagi tako dobljenega FSGP modela želimo načrtovati kompenzacijski regulator s spremenljivimi parametri, odvisnimi od spremenljivih parametrov FSGP modela. Potrebno je pripomniti, da bi lahko izbrali tudi eno od drugih metod načrtovanja vodenja. Na sliki 5.16 je prikazana shema regulacije za sistem drugega reda.

Ničle regulatorja izberemo tako, da so v vsakem trenutku  $t$  enake vrednostim polov reguliranega sistema. Tak način načrtovanja je uporaben samo za sisteme, ki so stabilni in fazno minimalni. Če to niso, je potrebno izbrati drugo metodo načrtovanja lokalnih regulatorjev. Za izbrani postopek načrtovanja potrebujemo spremenljive vrednosti polov, ki so določene z napovedanimi vrednostmi parametrov  $\hat{a}_1(\rho)$  in  $\hat{a}_2(\rho)$  in jih dobimo iz FSGP modela. Ojačenje regulatorja  $K_r$  skupaj z napovedanima vrednostima parametrov  $\hat{b}_1(\rho)$  in  $\hat{b}_2(\rho)$  določa zaprtozančno ojačenje sistema  $K_{cl}$ . Ker je FSGP model dobro identificiran le v področju blizu ravnotežne krivulje, pazimo, da ojačenja ne nastavimo previsoko. Zaradi razmeroma počasnih sprememb vhodnega signala tako sistem ostaja na zadovoljivo opisanem področju. Ojačenju regulatorja  $K_r$  smo s simulacijo s FSGP modelom določili vrednost  $K_r = 0.4$ . Diferenčna enačba, ki opisuje tako načrtan regulator, je:

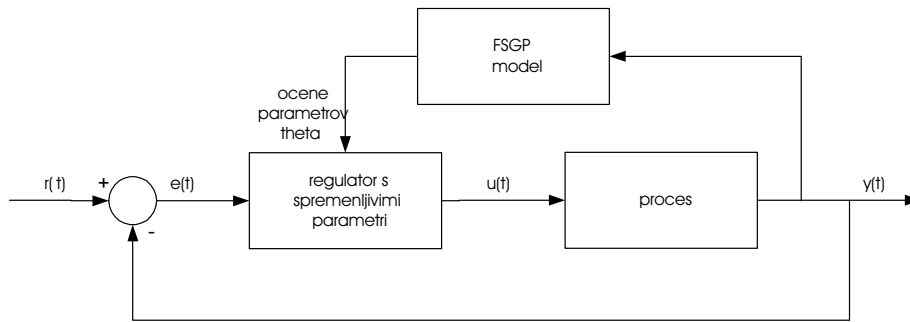
$$u(t) = u(t - T) + q_0(\rho) e(t) + q_1(\rho) e(t - T) + q_2(\rho) e(t - 2T), \quad (5.2)$$

**Tabela 5.5:** Hiperparametri GP modelov, ki opisujejo posamezne parametre lokalnih modelov sistema dveh posod

	$a_1(\rho)$	$a_2(\rho)$	$b_1(\rho)$	$b_2(\rho)$
$w_1$	$9.8 \cdot 10^{-5}$	$2.1 \cdot 10^{-4}$	$2.1 \cdot 10^{-10}$	$2.5 \cdot 10^{-4}$
$v$	0.85	0.080	1.40	0.60
$v_0$	$1.14 \cdot 10^{-3}$	$7.9 \cdot 10^{-4}$	$2.6 \cdot 10^{-2}$	$1.27 \cdot 10^{-2}$



**Slika 5.15:** Primerjava izhoda modelne naprave z izhodom FSGP modela med simulacijo (zgoraj), vrednosti parametrov FSGP modela med simulacijo skupaj z ustreznimi merami negotovosti  $2\sigma_{\theta_i}$  (spodaj)



**Slika 5.16:** Regulacijska zanka za sistem dveh posod

pri čemer so vrednosti spremenljajočih se parametrov  $q_0(\rho)$ ,  $q_1(\rho)$  in  $q_2(\rho)$  odvisni od napovedanih vrednosti parametrov  $\hat{a}_1(\rho)$ ,  $\hat{a}_2(\rho)$ ,  $\hat{b}_1(\rho)$  in  $\hat{b}_2(\rho)$ :

$$q_0(\rho) = 0.4 \frac{1}{\hat{b}_1(\rho) + \hat{b}_2(\rho)} \quad (5.3)$$

$$q_1(\rho) = q_0(\rho) \hat{a}_1(\rho) \quad (5.4)$$

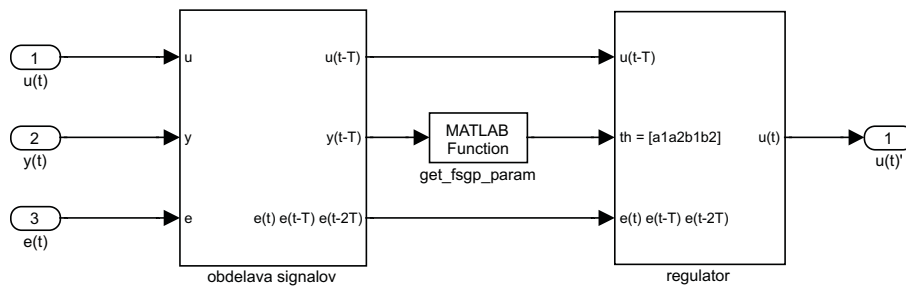
$$q_2(\rho) = q_0(\rho) \hat{a}_2(\rho) \quad (5.5)$$

Podatke o zaupanju v parametre modela  $\sigma_{\theta_i}$ ,  $i = 1, \dots, 4$ , bi lahko uporabili pri načrtovanju vodenja na več načinov: za načrtovanje robustnega vodenja, npr. za izračun optimalnih vrednosti parametrov regulatorja glede na odstopanja, izračun optimalne vrednosti regulirne veličine, tj. minimalno variančno vodenje, itd. Reševanje teh problemov bi preseglo okvir disertacije.

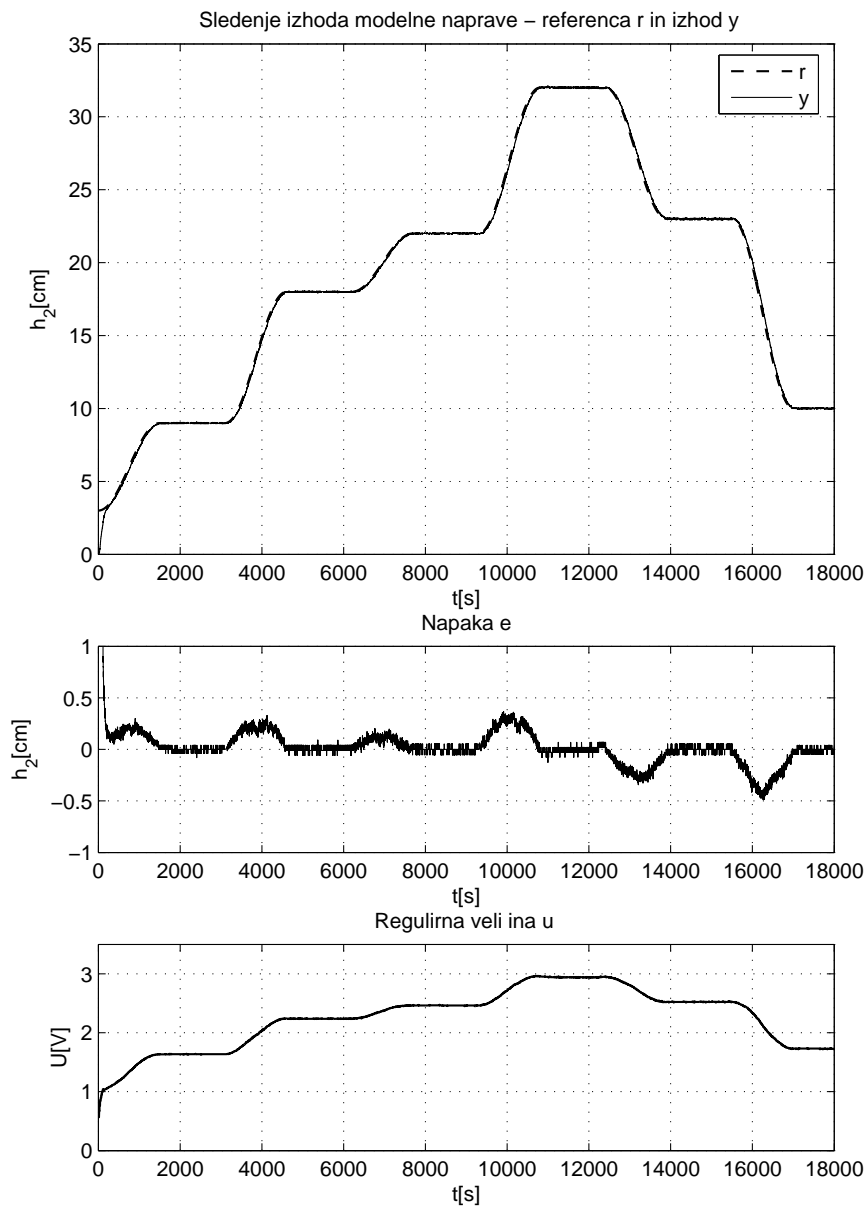
#### *Izvedba vodenja in rezultati*

Tudi pri izvedbi vodenja na napravi potrebujemo vrednosti spremenljivih parametrov modela, odvisnih od delovnega območja, saj določajo parametre regulatorja. Pomagamo si s funkcijo, opisano v Dodatku C.4. Shemo izvedbe regulacijske zanke v okolju Matlab/Simulink si lahko ogledamo na sliki 5.17.

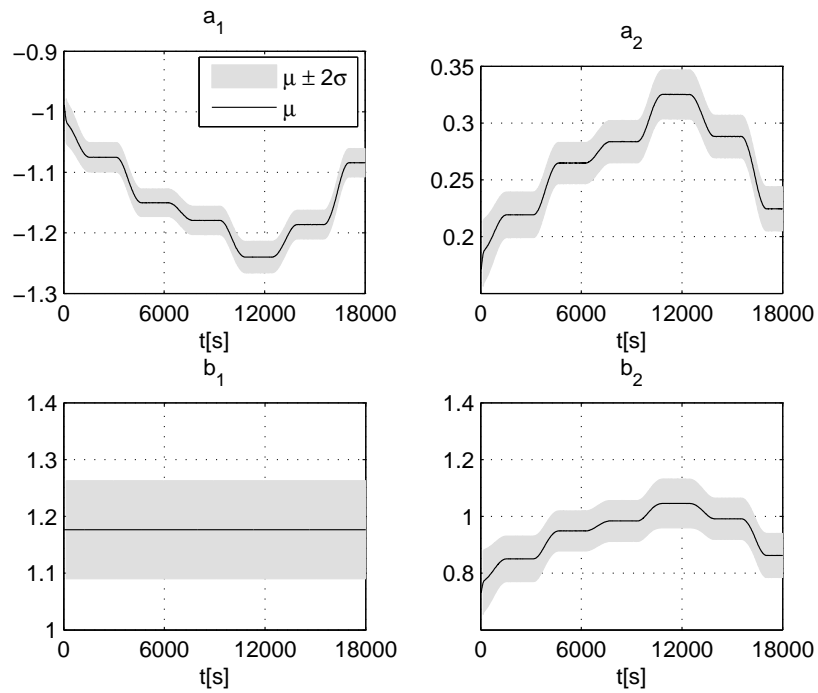
Rezultate tako izvedenega vodenja na modelni napravi si lahko ogledamo na sliki 5.18, vrednosti spremenljajočih se parametrov pa na sliki 5.19. Lahko opazimo, da izhod naprave lepo sledi referenci, ki je izbrana tako, da zaprtozančni sistem ne gre bistveno stran od ravnotežja. Če bi to želeli, bi morali vključiti tudi informacijo o neravnotežnih lokalnih modelih.



**Slika 5.17:** Izvedba regulacije v programskem paketu Matlab/Simulink za sistem dveh posod na podlagi FSGP modela. Signal  $u(t)$  vstopa v proces prek D/A pretvornika z dodanim zadrževalnikom ničtega reda (ZOH) s časom vzorčenja  $T_s = 10$  s. Prav tako je s časom vzorčenja  $T_s = 10$  s vzorčen merjeni signal  $y(t)$ .



**Slika 5.18:** Vodenje modelne naprave na podlagi FSGP modela. Razlika med izhodom naprave  $y$  in referenco  $r$  je zaradi dobrega sledenja težko opazna.



**Slika 5.19:** Spreminjajoče se napovedi parametrov skupaj z dvakratno standardno deviacijo med vodenjem modelne naprave

## 5.5 Identifikacija procesa priprave plina z uporabo FSGP modela za načrtovanje vodenja

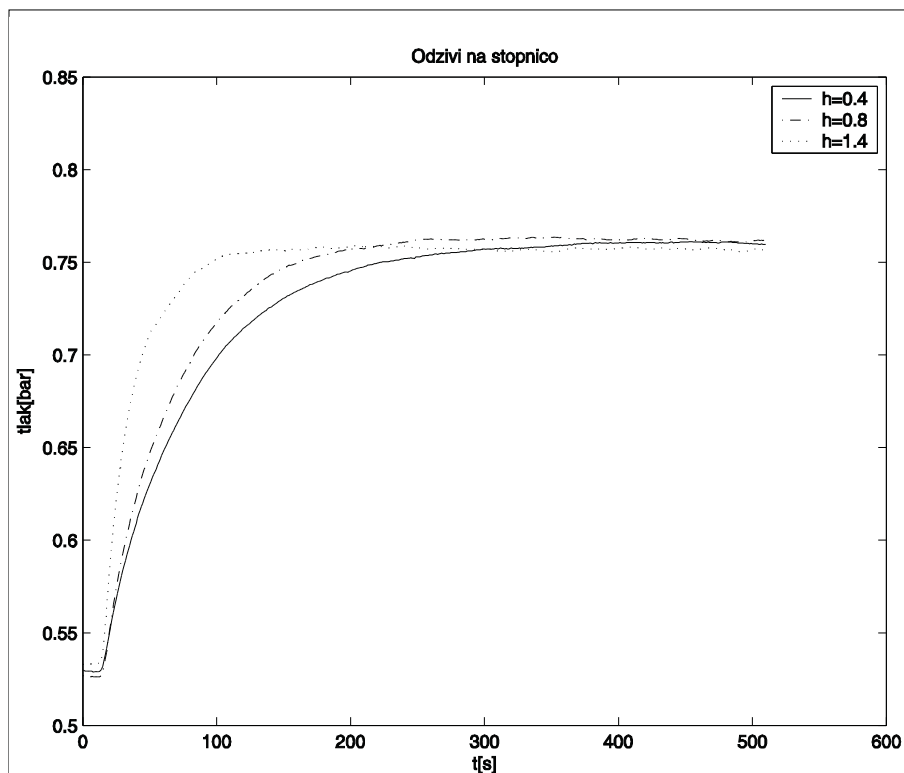
Želimo identificirati model procesa priprave plina [61, 62], predstavljenega v Dodatku E.2, na podlagi katerega bomo tako kot v prejšnjem podpoglavju, načrtovali preprosto vodenje s kompenzacijskim regulatorjem. Ker za načrtovanje takega regulatorja potrebujemo parametrični model, se odločimo za FSGP model. Poudarek tega primera ni na metodi načrtovanja vodenja, ampak na metodi modeliranja procesa.

V nadaljevanju si bomo ogledali identifikacijo sistema s FSGP modelom in njegovo vrednotenje tudi z uporabo za načrtovanje vodenja procesa. Načrtovali bomo zaprtozančni sistem, ki bo deloval v bližini ravnotežne krivulje procesa, saj je delovanje procesa, kot ga bomo predstavili, v glavnem omejeno le na to področje.

### Identifikacija procesa s FSGP modelom

Proces priprave plina (slika E.3), tudi ločevalnik plina in tekočine, je multivariabilen nelinearni proces, v katerem kot regulirani veličini  $y_1$  in  $y_2$  nastopata tlak  $p$  na področju od 0.4 do 0.7 bar in nivo vode  $h_1$  v tlačni posodi R4.1 na področju od 0.4 do 1.5 m. Regulirna signala sta  $u_1$  za odprtost ventila V4101 na izhodu plina iz posode in  $u_2$  za odprtost ventila V4102 na izhodu tekočine iz posode, oba na področjih od 0 (zaprt ventil) do 1 (odprt ventil). Glavni viri nelinearnosti procesa so nelinearne odvisnosti pretokov skozi ventile od tlačnih razlik na ventilih in nelinearna odvisnost dinamike tlaka plina od nivoja vode v tlačni posodi. Proces je multivariabilen, z dvema vhodoma in dvema izhodoma, z dinamičnimi križnimi povezavami. Relaciji med vhodoma  $u_1, u_2$  in nivojem  $h_1$  sta integrirnega značaja. Naša naloga je regulacija tlaka  $p$ , zato bomo proces obravnavali kot univariabilen nelinearni sistem z vhodno regulirno veličino  $u_1$  (odprtost ventila V4101), nivo  $h$ , reguliran s PI regulacijsko zanko, pa bomo upoštevali kot enega izmed merljivih dejavnikov nelinearnosti. Vplive navzkrižnih povezav v procesu bomo obravnavali kot nemerljive motnje. Glede na namen modeliranja procesa bomo v nadaljevanju tlak  $p$  označevali tudi z  $y$ , nivo tekočine s  $h$ , odprtost ventila  $u_1$  pa z  $u$ .

Iz matematičnega opisa procesa, enačbe (E.2), vidimo, da je sistem nelinearen. Na nelinearnost procesa lahko sklepamo tudi iz primerjav odprtozančnih odzivov tlaka  $p$  na stopničasto spremembo položaja ventila V4101 od  $u_1 = 0.51$  do  $u_1 = 0.45$  pri različnih vrednostih nivoja  $h$ , prikazanih na sliki 5.20.



**Slika 5.20:** Odprtozančni odziv tlaka  $p$  na stopničasto spremembo položaja ventila V4101 od  $u_1 = 0.51$  do  $u_1 = 0.45$  pri različnih vrednostih nivoja  $h$  [61]

Na posameznih območjih delovanja, ta so določena z nivojem tekočine  $h$ , bomo identificirali linearne lokalne modele prvega reda oblike:

$$y(t + T) = -a_1(t)y(t) + b_2(t)u(t - T). \quad (5.6)$$

z dvema spremenljivima parametroma  $a_1(t)$  in  $b_2(t)$ . Za model prvega reda smo se odločili, ker identifikacija z višjim redom ni prinesla izboljšanja rezultatov. To je v skladu z ugotovitvami drugih, ki so modelirali ta proces [62].

Odvisnost odziva sistema od višine tekočine  $h$  upoštevamo tako, da višino tekočine  $h$  izberemo za razvrstilno spremenljivko  $\rho(t) = h(t)$ .

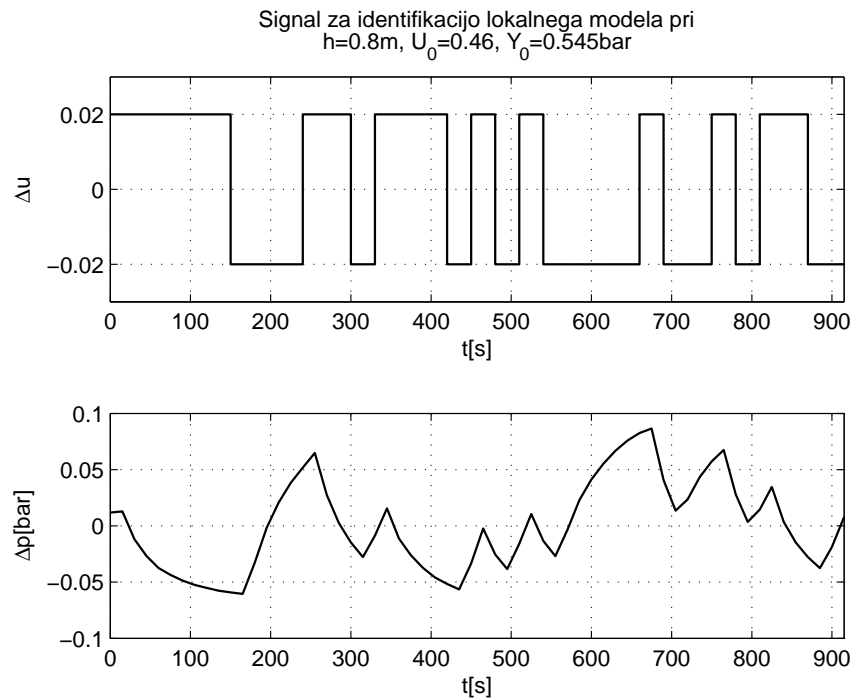
#### *Identifikacija parametrov lokalnih modelov*

Za identifikacijo lokalnih modelov s parametroma  $a_1$  in  $b_2$ , odvisnima od stanja procesa, ki je določen z višino tekočine  $h$ , potrebujemo tri signale: položaj ventila V4101  $u$ , nivo tekočine  $v$  in tlak  $y$ . Pred identifikacijo posameznega lokalnega modela smo sistem najprej pripeljali v zeleno območje delovanja z regulacijo

nivoja tekočine  $h$ . Ko se je nivo ustalil na želeni vrednosti, smo odprtosti ventila  $U_0 = 0.46$  dodali PRBS signal z amplitudo  $\Delta u = 0.02$  in opazovali vpliv na izhodu  $y$ .

Linearni lokalni modeli so bili identificirani z metodo najmanjših kvadratov s posplošenim pogreškom. Na sliki 5.21 sta prikazana vhodni in izhodni signal za identifikacijo pri nivoju  $h = 0.8$  m, na sliki 5.22 pa je za isto delovno točko prikazana primerjava odziva procesa in identificiranega lokalnega modela. Iz slednje lahko razberemo za naš namen ustrezno ujemanje dinamike.

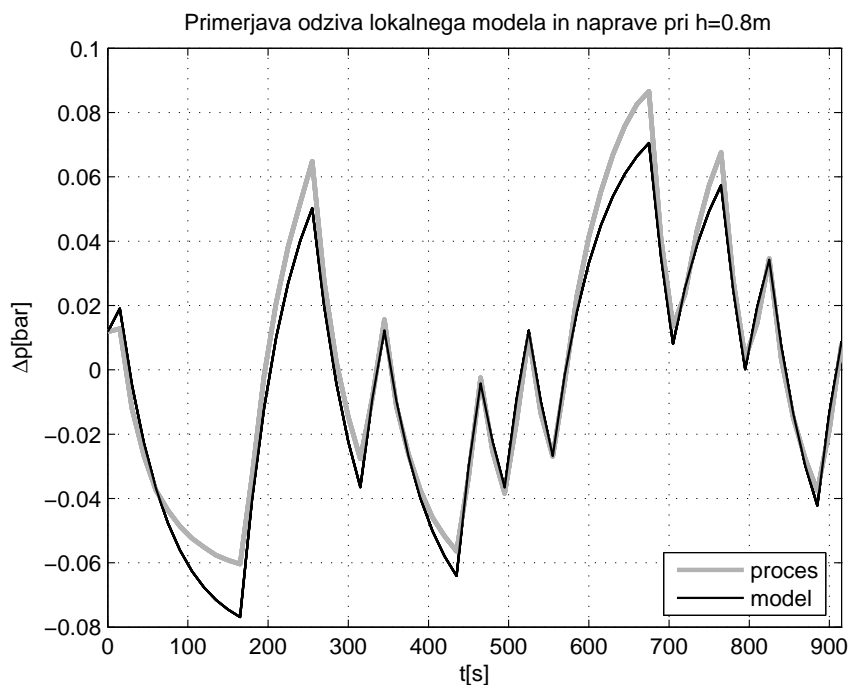
S tem postopkom smo pri času vzorčenja  $T_s = 1$  s identificirali 11 lokalnih modelov pri različnih (posejanih enakomerno med  $h = 0.4$  m in  $h = 1.4$  m) nivojih tekočine. Vrednosti parametrov pri ustreznih vrednostih razvrstilne spremenljivke  $h$  so skupaj s svojimi variancami zbrane v tabeli 5.6.



**Slika 5.21:** Signal za identifikacijo lokalnega modela pri  $h = 0.8$  m

### *Tvorjenje in vrednotenje FSGP modela*

FSGP model procesa sestavljata dva GP modela, ki opisujeta vrednosti parametrov  $a_1$  in  $b_2$  v odvisnosti od razvrstilne spremenljivke  $h(t)$ . Za kovariančno funkcijo smo zaradi predpostavke o gladkosti spreminjanja parametrov lokalnih modelov ponovno izbrali kar Gaussovo kovariančno funkcijo (2.4).



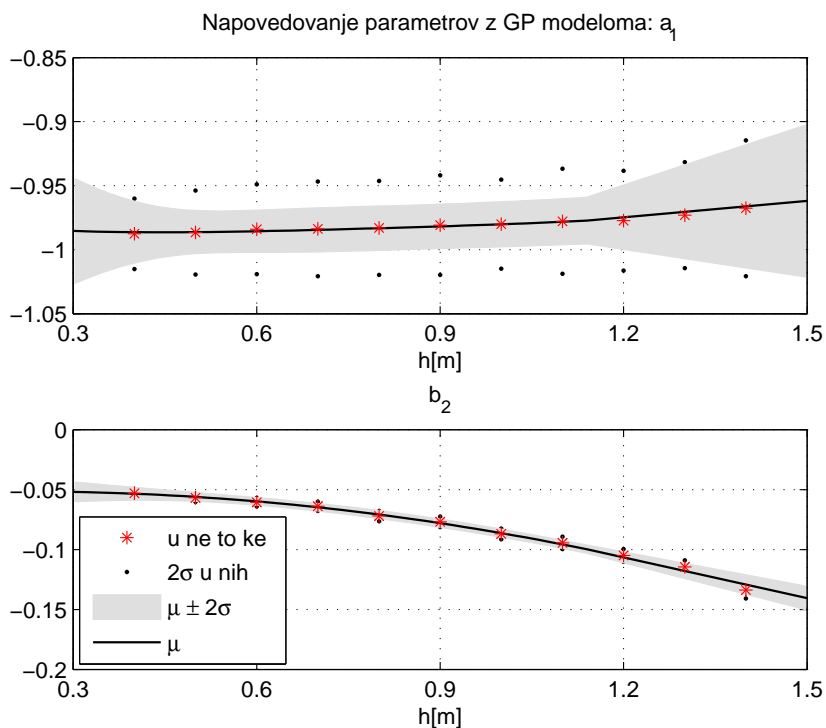
**Slika 5.22:** Primerjava odziva lokalnega modela in naprave pri  $h = 0.8$  m

Parametri lokalnih modelov in ustrezne vrednosti razvrstilne spremenljivke  $h(t)$  tako predstavljajo učne podatke za učenje GP modelov. Za učenje hiperparametrov smo, kot v celotni disertaciji, uporabili metodo največje podobnosti.

**Tabela 5.6:** Parametri identificiranih lokalnih modelov procesa skupaj z ustreznimi variancami.

$h$	$a_1$	$\text{var}(a_1)$	$b_2$	$\text{var}(b_2)$
0.4	-0.988	$1.89 \cdot 10^{-4}$	-0.053	$1.99 \cdot 10^{-6}$
0.5	-0.987	$2.68 \cdot 10^{-4}$	-0.057	$2.86 \cdot 10^{-6}$
0.6	-0.984	$3.07 \cdot 10^{-4}$	-0.061	$3.16 \cdot 10^{-6}$
0.7	-0.984	$3.41 \cdot 10^{-4}$	-0.064	$3.67 \cdot 10^{-6}$
0.8	-0.983	$3.36 \cdot 10^{-4}$	-0.072	$4.70 \cdot 10^{-6}$
0.9	-0.981	$3.79 \cdot 10^{-4}$	-0.077	$4.82 \cdot 10^{-6}$
1.0	-0.980	$3.02 \cdot 10^{-4}$	-0.087	$4.95 \cdot 10^{-6}$
1.1	-0.978	$4.22 \cdot 10^{-4}$	-0.094	$6.71 \cdot 10^{-6}$
1.2	-0.977	$3.79 \cdot 10^{-4}$	-0.105	$7.12 \cdot 10^{-6}$
1.3	-0.973	$4.28 \cdot 10^{-4}$	-0.114	$7.34 \cdot 10^{-6}$
1.4	-0.968	$7.03 \cdot 10^{-4}$	-0.134	$1.26 \cdot 10^{-5}$

Rezultati učenja, pri katerem varianc identificiranih parametrov lokalnih modelov nismo vključevali, so dali zelo majhne vrednosti hiperparametra  $v_0$ , ki prispeva k velikosti napovedane variance parametrov  $a_1$  in  $b_2$ . Tako majhne vrednosti lahko pojasnimo, če si na sliki 5.23 ogledamo učne podatke, označene z \*. Lahko razberemo, da je napovedovanje parametrov v odvisnosti od razvrstilne spremenljivke  $h$  enostavno, kar pojasnjuje majhnost ocenjene velikosti šuma. Zato smo se odločili, da v učne podatke vključimo tudi ocenjene variance identificiranih parametrov in GP modela še enkrat učili. Vrednosti posameznih hiperparametrov GP modelov po tem učenju so zbrane v tabeli 5.7, napovedovanje parametrov, skupaj s tako dobljenimi merami negotovosti in z učnimi podatki, pa si lahko ogledamo na sliki 5.23.



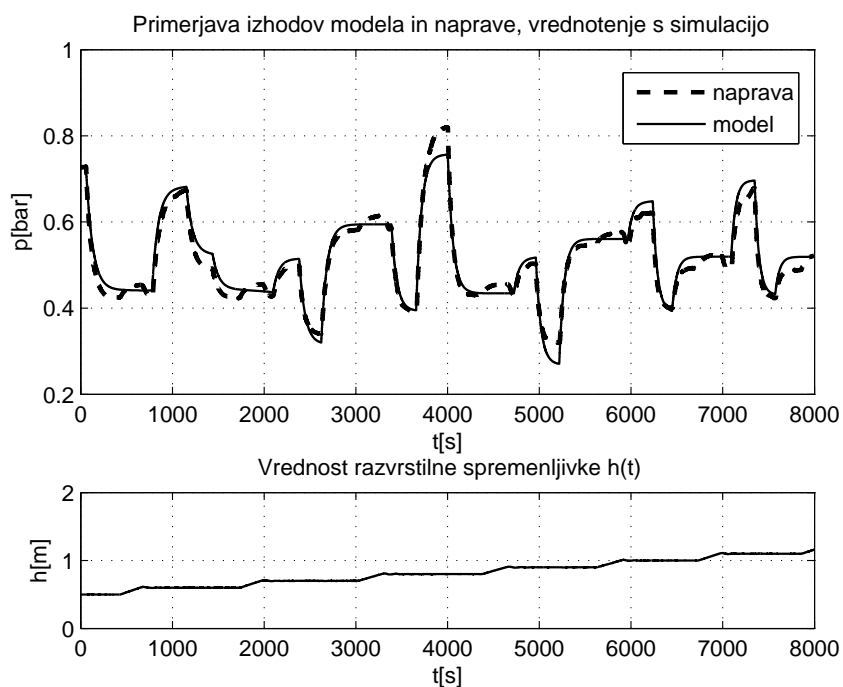
**Slika 5.23:** Napovedovanje parametrov modela z GP modeli skupaj z učnimi podatki. Z zvezdicami so označeni identificirani parametri lokalnih modelov (učne točke), s pikami pa njihova dvakratna standardna deviacija, ki smo jo tudi uporabili za učenje GP modelov.

Naučena GP modela smo združili v FSGP model, kot je bilo predstavljeno v podpoglavju 4.3. Za vrednotenje FSGP modela smo uporabili simulacijo s podobnimi signali, kot smo jih pozneje uporabili za vodenje. Algoritem uporabljene funkcije za simulacijo FSGP modela je predstavljen v Dodatku C.3. Primerjava izhoda modelne naprave in odziva FSGP modela na signalu za vrednotenje je, skupaj z vrednostmi razvrstilne spremenljivke  $h(t)$ , prikazana

**Tabela 5.7:** Hiperparametri GP modelov, ki opisujejo posamezne parametre lokalnih modelov procesa

	$a_1(\rho)$	$b_2(\rho)$
$v$	0.95	0.03
$w_1$	0.60	0.45

na sliki 5.24. Opazno je rahlo odstopanje vrednosti izhoda modela in naprave pri konstantnem vходу v različnih delovnih točkah. Vzrok za to bi bil lahko vpliv mehanskih izvršilnih sistemov. Ta napaka nas ne moti, saj jo bo odpravil načrtovani regulator. Vrednosti spremenljivih parametrov FSGP modela med simulacijo so prikazane na sliki 5.25.

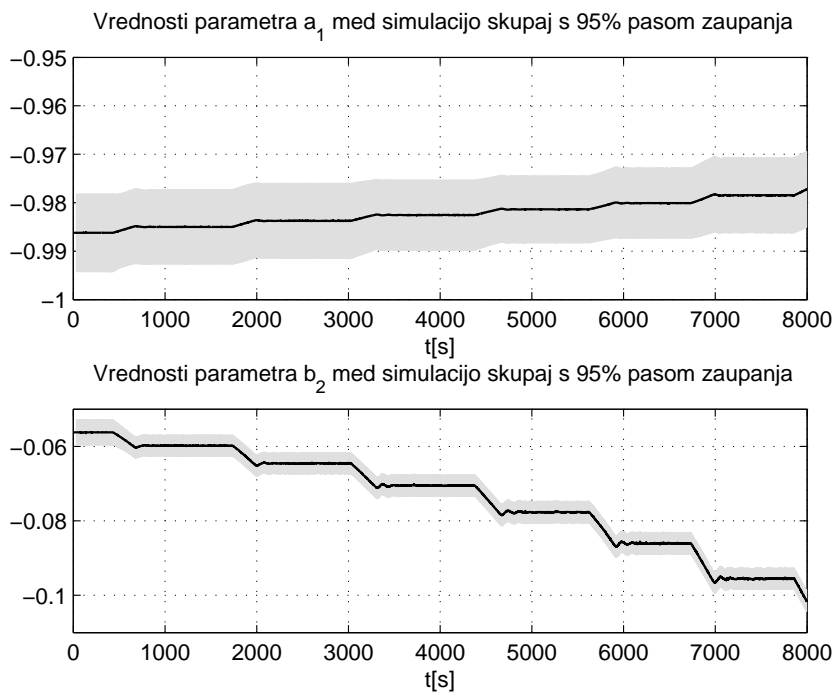


**Slika 5.24:** Primerjava izhoda modelne naprave in izhoda FSGP modela med simulacijo, skupaj z razvrstilno spremenljivko  $h$ .

## Vodenje procesa priprave plina z uporabo FSGP modela

### *Načrtovanje vodenja z uporabo FSGP modela*

Na podlagi tako dobljenega FSGP modela smo načrtovali kompenzacijski regulator s spremenljivimi parametri, odvisnimi od ocen spremenljivih



**Slika 5.25:** Vrednosti parametrov FSGP modela med simulacijo skupaj s 95% pasom zaupanja

parametrov  $\hat{a}_1$  in  $\hat{b}_2$ , dobljenimi z uporabo FSGP modela. Velja pripomniti, da bi lahko izbrali tudi eno izmed drugih metod načrtovanja vodenja.

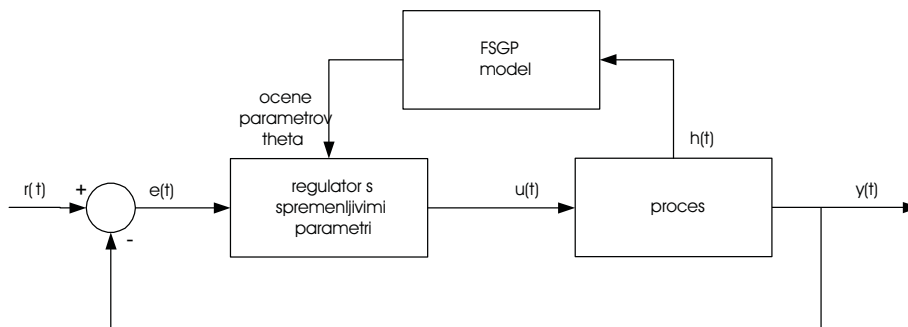
Shema regulacije je prikazana na sliki 5.26. Ničlo regulatorja izberemo tako, da je enaka kot vrednost pola reguliranega sistema. Kot smo omenili v prejšnjem podpoglavju, je tak način načrtovanja uporaben samo za stabilne in fazno minimalne sisteme, sicer je potrebno izbrati drugo metodo načrtovanja lokalnih regulatorjev. Za izbrani postopek načrtovanja potrebujemo spremenljivo vrednost pola, ki jo dobimo iz napovedi parametra  $\hat{a}_1(h(t))$  iz FSGP modela. Ojačenje regulatorja  $K_r$ , ki skupaj z oceno spremenljivega parametra  $\hat{b}_2$  določa zaprtozančno ojačenje sistema, izberemo tako, da ojačenje v regulacijski zanki ni previsoko. Tako zagotovimo, da proces ostaja v bližini ravnotežne krivulje, kjer je s FSGP modelom razmeroma dobro opisan. S simulacijo zaprtozančnega sistema s FSGP modelom in nato preizkusom na napravi smo ojačenje regulatorja določili kot  $K_r = 0.1$ . Diferenčna enačba, ki opisuje tako načrtan regulator, je:

$$u(t) = u(t - T) + q_1(\rho) e(t - T) + q_2(\rho) e(t - 2T), \quad (5.7)$$

pri čemer sta vrednosti spreminjajočih se parametrov  $q_1$  in  $q_2$  odvisni od napovedanih vrednosti parametrov  $\hat{a}_1(\rho)$  in  $\hat{b}_2(\rho)$ :

$$q_1(\rho) = \frac{0.1}{\hat{b}_2(\rho)} \quad (5.8)$$

$$q_2(\rho) = q_1(\rho) \hat{a}_1(\rho) \quad (5.9)$$

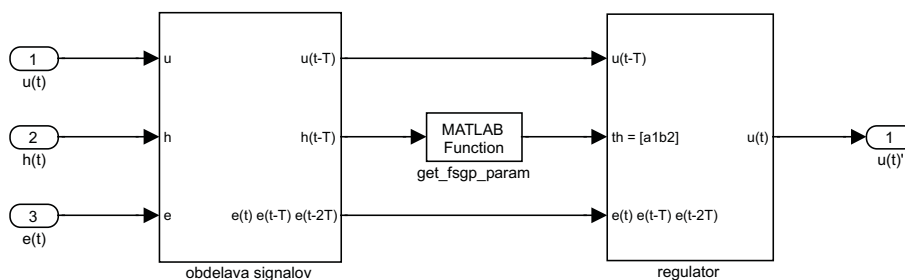


**Slika 5.26:** Regulacijska zanka za proces priprave plina

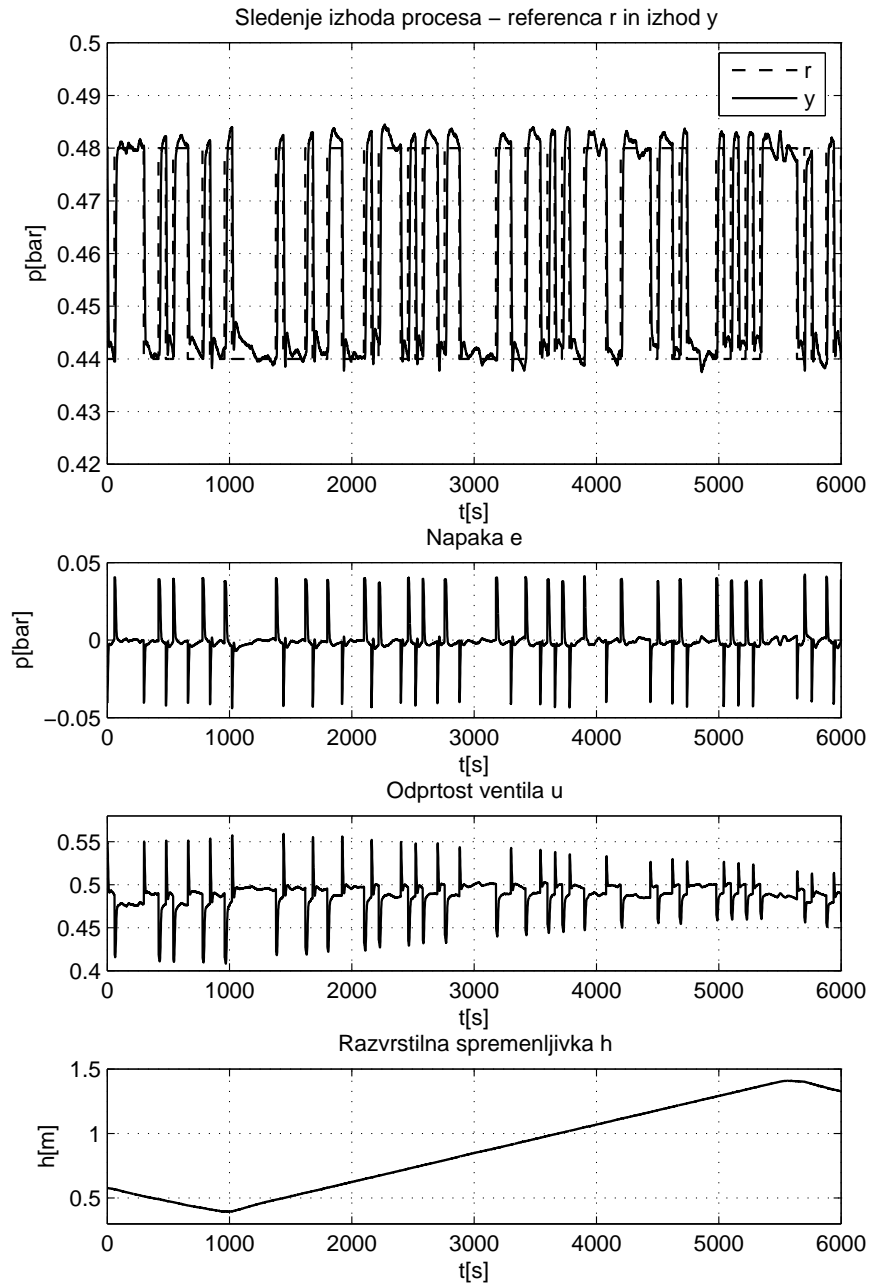
### *Izvedba vodenja in rezultati*

Pri izvedbi vodenja na napravi potrebujemo vrednosti spremenljivih parametrov modela, odvisnih od delovnega območja, saj določajo parametre regulatorja. Pomagamo si s funkcijo, opisano v Dodatku C.4. Shemo izvedbe regulacijske zanke v okolju Matlab/Simulink si lahko ogledamo na sliki 5.27.

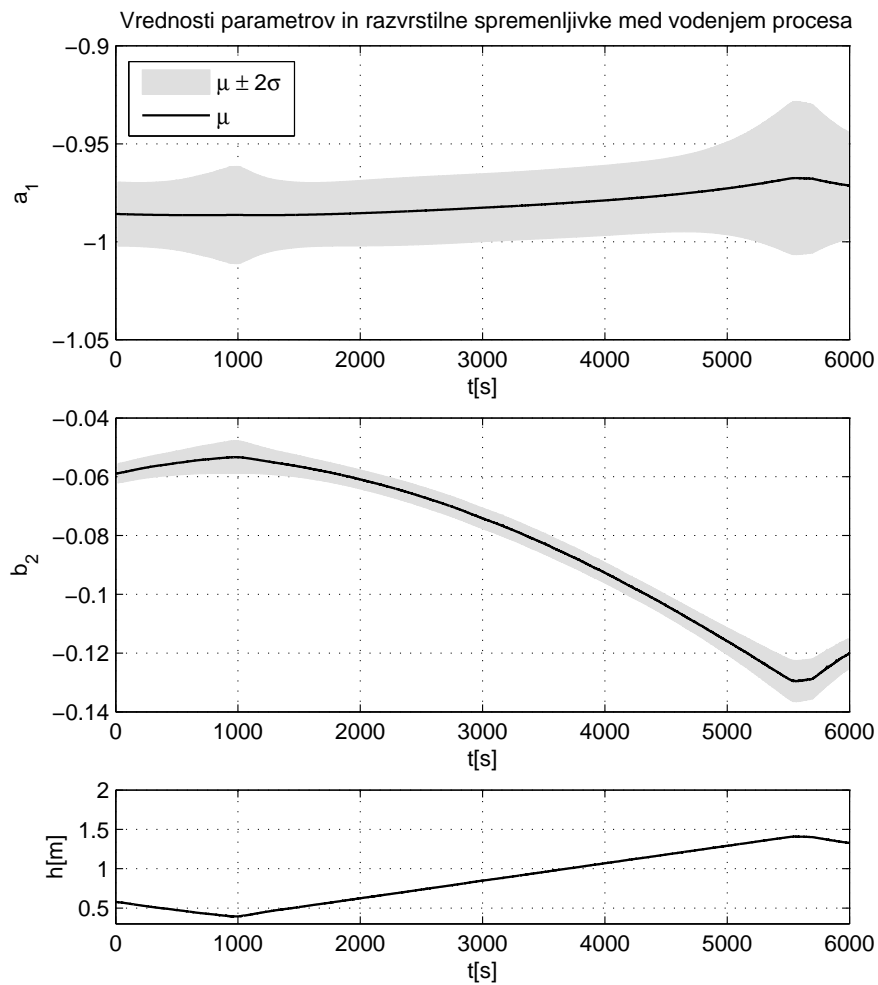
Rezultate tako izvedenega vodenja na modelni napravi si lahko ogledamo na sliki 5.28, ustrezne vrednosti spreminjajočih se parametrov pa na sliki 5.29. Izhod procesa sledi referenci po celotnem področju ne glede na spreminjanje dinamike procesa v odvisnosti od višine tekočine v napravi. Vodenje bi lahko izboljšali, vendar to ni bil namen predstavljenega primera.



**Slika 5.27:** Izvedba regulacije v okolju Matlab/Simulink za sistem dveh posod na podlagi FSGP modela



Slika 5.28: Vodenje procesa na podlagi FSGP modela



**Slika 5.29:** Spreminjajoče se napovedi parametrov  $a_1$  in  $b_2$  skupaj z dvakratno standardno deviacijo med vodenjem procesa

## 5.6 Identifikacija sistema dveh posod na hidravlični modelni napravi za zaznavanje napak

Vodenje in nadzor procesov v industriji sta vse bolj avtomatizirana, kar ima seveda svoje prednosti (boljša kakovost, večja varnost, nižja cena izdelkov itd.). Problem se pojavi, če v nekem procesu pride do napake. V takem primeru bi želeli napako na procesu odkriti čim prej, da zmanjšamo ekološko in ekonomsko škodo in morda preprečimo zdravju ali življenju nevarne razmere.

V ta namen za nadzor procesov uporabljamo različne vrste zaznavanja napak (angl. *fault detection*, FD) [25]. Nadzor procesov lahko opravlja človek, vendar v današnji procesni industriji zelo veliko razpoložljivih informacij lahko povzroči izgubo pregleda nad situacijo, zato je ugodno, če vsaj del nadzora opravljajo računalniki. Metode zaznavanja napak delimo na tri družine. Kvantitativne metode za zaznavanje napak uporabljajo identificirane modele sistema [111], kvalitativne metode za zaznavanje napak uporabljajo znanje o osnovnih principih (fizikalnih, kemijskih itd.), na katerih temelji obnašanje sistema [112], metode zaznavanja napak na podlagi zgodovine procesa pa uporabljajo izključno veliko število že znanih in primerno obdelanih podatkov o procesu [113]. Nas bo zanimalo področje zaznavanja napak, kjer za FD uporabljamo identificirane modele. To je v primeru, da je sistem/proces opisan z linearnimi modeli, dobro opisano, npr. [25, 111], v primeru opisa sistema/procesa z nelinearnim modelom pa je problem zaznavanja napak večji [41].

Po tem pristopu za zaznavanje napak na procesih največkrat primerjamo izhod procesa in izhod njegovega modela (observatorja) oz. opazujemo njuno razliko, ki ji pravimo *residual*. V primeru enakih izhodov (*residual* je enak nič) domnevamo, da proces deluje, kot je predvideno, če pa je *residual* različen od nič, domnevamo, da je v procesu prišlo do napake. Pri vsem tem domnevamo, da model procesa dovolj dobro opisuje proces na vseh delovnih področjih. Do problema pride, ko model ne daje natančne informacije o procesu. V takem primeru lahko nepopolen ali netočen opis procesa vodi do različnih izhodov procesa in observatorja in posledično do lažnega alarma. To je stanje, ko algoritem za zaznavanje napak javi napako na procesu, čeprav ta deluje normalno. Ena od možnosti rešitve tega problema je zvišanje praga (angl. *threshold*) residuala, pri katerem algoritem javi napako na procesu. Tako je mogoče zmanjšati ali eliminirati lažne alarme, toda hkrati zmanjšamo tudi občutljivost sistema za zaznavanje napak, saj je za sproženje alarma potrebna večja napaka. Druga možna rešitev problema je časovno spremenljivo nastavljanje praga, odvisno od stopnje zaupanja v model [22]. Še ena možnost povečanja robustnosti je modeliranje vpliva napak modeliranja sistema na residuele in njihova sprotna kompenzacija [42].

V tem podpoglavju bomo pokazali, kako si pri problemu slabega opisa sistema za zaznavanje napak na nekem področju pomagamo z uporabo GP oz. LMGP modela. Omejili se bomo na zaznavanje napak na senzorjih izhodne veličine. Glavna ideja postopka, predlaganega v [40], je izkoristiti informacijo o kvaliteti modela na določenem področju delovanja, dano z mero zaupanja v napoved GP modela. Kvaliteto opisa sistema bomo ocenili z veljavnostnim indeksom (angl. *validity index*), za zaznavanje napak pa bomo uporabili test verjetnosti hipotez. Zaznavanje napak bomo prikazali na primeru sistema dveh posod laboratorijske modelne naprave, opisanega z LMGP modelom.

## Zaznavanje napak z uporabo GP modela

Za ocenjevanje pristranskosti senzorjev bomo uporabljali primerjavo izhodov sistema  $y$  in observatorja  $\hat{y}$  oz. njune razlike, tj. residuala:

$$e(k) = y(k) - \hat{y}(k). \quad (5.10)$$

Če za napovedovanje izhoda sistema uporabljamo GP model, ima v koraku  $k$  napovedan izhod obliko Gaussove porazdelitve:  $\hat{y}(k) \sim \mathcal{N}(m(k), v(k))$  in residual  $e(k)$  nosi informacijo tako o najbolj verjetni vrednosti napake napovedi  $\varepsilon(k) = y(k) - m(\mathbf{x}(k))$  kot o zaupanju v napoved  $v(\mathbf{x}(k))$ , kjer smo z  $\mathbf{x}(k)$  označili vhod v GP model v  $k$ -tem koraku.

Intuitivno si sistem zaznavanja lahko zamislimo takole:

- Če je razlika med najbolj verjetno napovedjo izhoda modela in izhodom sistema  $\varepsilon(k)$  majhna, potem do napake na sistemu verjetno ni prišlo.
- Če je razlika med najbolj verjetno napovedjo modela in izhodom sistema  $\varepsilon(k)$  velika, potem je:
  - a. prišlo do napake pri merjenju izhoda sistema ali pa
  - b. je sistem na področju, kjer je z modelom slabo opisan. V primeru GP modela slab opis pomeni, da je vektor regresorjev  $\mathbf{x}(k)$ , ki nastopa kot vhod v GP model, daleč stran od katerega izmed vhodnih vektorjev  $\mathbf{x}_i$  iz učne množice  $\mathcal{D}$ , na kar nas GP model opozori s povečano varianco napovedi  $v(\mathbf{x}(k))$ .

Na podlagi opazovanja najbolj verjetne vrednosti napake napovedi  $\varepsilon(k)$  in variance napovedi  $v(\mathbf{x}(k))$  se tako lahko odločimo za najbolj verjetno izmed treh naštetih možnosti.

Najbolj enostavno bi bilo zaznavanje napak na podlagi samo ene (zadnje) vrednosti residuala, vendar ta možnost v praksi ni najbolj priporočljiva, saj se poveča vpliv šuma na zaznavanje, kar vodi do slabše *stabilnosti zaznavanja* (angl. *detection stability*). Za detektor pravimo, da je nestabilen, če pogosto prehaja med stanjema “je napaka” – “ni napake”. Stabilnost zaznavanja lahko izboljšamo, če uporabimo več residualov naenkrat, npr. okno  $N$  zadnjih residualov [70]. Večje, kot je okno, stabilnejši, a hkrati tudi počasnejši, je detektor.

S tem v mislih definiramo vektor residualov  $\mathbf{e}(k)$  kot:

$$\mathbf{e}(k) = [e(k - N + 1) \ e(k - N + 2) \ \dots \ e(k)], \quad (5.11)$$

ki ga lahko predstavimo z vektorjem srednjih vrednosti residualov  $\boldsymbol{\varepsilon}(k)$  in ustrezno kovariančno matriko  $\boldsymbol{\Sigma}(k)$  kot:

$$\boldsymbol{\varepsilon}(k) = [\varepsilon(k - N + 1) \ \dots \ \varepsilon(k)], \quad (5.12)$$

$$\boldsymbol{\Sigma}(k) = \text{diag} [v(\mathbf{x}(k - N + 1)) \ \dots \ v(\mathbf{x}(k))]. \quad (5.13)$$

Naša naloga je, da se glede na vektor residualov  $\boldsymbol{\varepsilon}(k)$  in kovariančno matriko  $\boldsymbol{\Sigma}(k)$  odločimo, ali je senzor izhodne veličine  $y(k)$  pristranski ( $f_y \neq 0$ ) ali ne ( $f_y = 0$ ). Pri tem  $f_y$  izraža pristranskost sensorja.

### Pravilo zaznavanja na podlagi statističnega testa

Če senzor izhoda ni pristranski, potem morajo srednje vrednosti napake  $\varepsilon(k)$  ležati blizu nič v vseh korakih  $k$ , njihova porazdelitev mora ustrezati predpostavljeni porazdelitvi šuma, pri čemer v našem primeru predpostavljamo beli Gaussov šum. Pristranskost sensorja izhoda ( $f_y \neq 0$ ) povzroči pristranskost residualov  $e(k)$ . Odločamo se med dvema možnostma:

1. ničelna hipoteza – senzor ni pristranski:  $H_0 : f_y = 0$  in
2. alternativna hipoteza – senzor je pristranski:  $H_1 : f_y \neq 0$ .

Med tema hipotezama se bomo odločili glede na razmerje med dvema verjetnostnima porazdelitvama [92, 115]. Hipotezo  $H_0$  zavrnamo oz. se odločimo za hipotezo  $H_1$ , če je razmerje med dvema verjetnostnima porazdelitvama:

$$\kappa = \frac{p_{H_0}(\boldsymbol{\varepsilon}(k))}{\sup_{f_y \neq 0} p_{H_1}(\boldsymbol{\varepsilon}(k))} < \tau. \quad (5.14)$$

Konstanto  $\tau$  določimo glede na to, kako veliko napako še dopuščamo, da bi povečali robustnost.

V izrazu (5.14) nastopa verjetnost hipoteze  $H_1$ . Hipoteza  $H_1$  je izmed vseh hipotez, ki predpostavljajo pristranskost senzorjev izhoda, tista, ki je pri danem vektorju najbolj verjetnih vrednosti residualov  $\boldsymbol{\varepsilon}(k)$  in kovariančni matriki  $\boldsymbol{\Sigma}(k)$  najbolj verjetna. Da jo določimo, predpostavljamo Gaussovo porazdelitev izhoda s srednjo vrednostjo  $\nu$  in kovariančno matriko  $\boldsymbol{\Sigma}$  [40]:

$$\begin{aligned} p(\boldsymbol{\varepsilon}(k), \nu) &= p(\boldsymbol{\varepsilon}(k - N + 1), \dots, \boldsymbol{\varepsilon}(k), \nu) \\ &= \left( \frac{1}{(2\pi)^N \prod v(\mathbf{x}(n))} \right)^{1/2} \exp \left[ -\frac{1}{2} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i}) \right], \end{aligned} \quad (5.15)$$

kjer je  $n = k - N + 1, \dots, k$  in je  $\mathbf{i}$  vektor enic dolžine  $N$ . Ekstrem porazdelitve  $p(\boldsymbol{\varepsilon}, \nu)$  pri predpostavljani obliki dobimo z njenim odvajanjem po  $\nu$  in izenačitvijo z nič:

$$\frac{\partial p(\boldsymbol{\varepsilon}(k), \nu)}{\partial \nu} = -\frac{1}{2} p(\boldsymbol{\varepsilon}, \nu) \frac{\partial ((\boldsymbol{\varepsilon}(k) - \nu \mathbf{i})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i}))}{\partial \nu} = 0. \quad (5.16)$$

Sledi:

$$\begin{aligned} \frac{\partial ((\boldsymbol{\varepsilon}(k) - \nu \mathbf{i})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i}))}{\partial \nu} &= 0 \\ -\mathbf{i}^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i}) + (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i})^T (-\mathbf{i}) &= 0 \\ -2 (\mathbf{i}^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i})) &= 0 \end{aligned}$$

in za najbolj verjetno pristranskost izhoda  $\nu$  dobimo:

$$\nu = \frac{\mathbf{i}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}(k)}{\mathbf{i}^T \boldsymbol{\Sigma}^{-1} \mathbf{i}}. \quad (5.17)$$

Preoblikujmo še izraz, ki določa mejo med hipotezama  $H_0$  in  $H_1$ . Z logaritmiranjem leve strani izraza (5.14) in uporabo pristranskosti  $\nu$  (5.17) dobimo:

$$\begin{aligned} \kappa &= \frac{p_{H_0}(\boldsymbol{\varepsilon}(k))}{\sup_{f_y \neq 0} p_{H_1}(\boldsymbol{\varepsilon}(k))} \\ &= \frac{p_{H_0}}{p_{H_1|\nu}} = \frac{\exp \left[ -\frac{1}{2} \boldsymbol{\varepsilon}(k)^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}(k) \right]}{\exp \left[ -\frac{1}{2} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i}) \right]} \\ \log \kappa &= \frac{1}{2} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\varepsilon}(k) - \nu \mathbf{i}) - \frac{1}{2} \boldsymbol{\varepsilon}(k)^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}(k) \\ -2 \log \kappa &= \frac{(\mathbf{i}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}(k))^2}{\mathbf{i}^T \boldsymbol{\Sigma}^{-1} \mathbf{i}}. \end{aligned} \quad (5.18)$$

Sedaj lahko pogoj (5.14) za zavrnitev hipoteze  $H_0$  napišemo kot [40]:

$$\begin{aligned} \sqrt{-2 \log \kappa} &> \sqrt{-2 \log \tau} \\ \frac{|\mathbf{i}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\varepsilon}(k)|}{\sqrt{\mathbf{i}^T \boldsymbol{\Sigma}^{-1} \mathbf{i}}} &> c_{1-\beta/2} = c_{\max}, \end{aligned} \quad (5.19)$$

kjer je  $c_{1-\beta/2}$  stopnja pomembnosti (angl. *level of significance*), ki pove, kolikšno mora biti razmerje verjetnosti porazdelitev hipotez  $H_1$  in  $H_0$ , da se odločimo za prvo.

### Veljavnostni indeks

Rezultati testa (5.19) so lahko zavajajoči, če je GP model primoran napovedovati izhod sistema na področju, ki ni opisano z vzorci v učni množici  $\mathcal{D}$ . Za boljšo interpretacijo rezultatov testa (5.19) uporabimo *veljavnostni indeks*  $I \in [0, 1]$ , [40], ki naj pove, koliko lahko rezultatu testa zaupamo. Vrednost indeksa  $I$  blizu ena bo pomenila, da testu lahko zaupamo. Vrednost indeksa  $I$  blizu nič pa bo pomenila, da je sistem na trenutnem delovnem območju slabo opisan in da moramo biti pri interpretaciji rezultatov testa (5.19) previdni, saj ne vemo, ali je prišlo do povečanja vrednosti residualov zaradi napake na sistemu ali samo zaradi slabšega opisa sistema.

Veljavnost  $I_k$  posameznega residuala  $\varepsilon(k)$  v koraku  $k$  določimo v odvisnosti od “razdalje”  $\delta_k$  vhodnega vektorja regresorjev  $\mathbf{x}(k)$  od točk, ki sestavljajo učno množico  $\mathcal{D}$ .

Ena izmed možnosti za določitev veljavnostnega indeksa  $I_k$  je opazovanje napovedane variance izhoda GP modela [40], v kateri je že zajet vpliv bližine učnih točk. Kot referenčno velikost variance izberemo največjo napovedano varianco pri napovedovanju izhoda v točkah učne množice, pri kateri pa ni upoštevana varianca šuma na izhodu (enačba 2.15 brez upoštevanja variance  $v_0$  v kovariančni matriki  $\mathbf{K}$ ):

$$\begin{aligned} \delta(\mathbf{x}_k, \mathbf{X}) &= k(\mathbf{x}_k) - \mathbf{k}(\mathbf{x}_k)^T \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}_k) \\ \delta_{\max} &= \max(\delta(\mathbf{x}_k, \mathbf{X})), k = 1, \dots, N, \end{aligned} \quad (5.20)$$

kjer so  $\mathbf{x}_k$  vhodni vektorji točk učne množice. Sedaj lahko veljavnostni indeks  $I_k$  v  $k$ -tem koraku določimo kot:

$$I_k = \begin{cases} 1 & \delta_{\max} \geq \delta(\mathbf{x}(k), \mathbf{X}) \\ \frac{\delta_{\max}}{\delta(\mathbf{x}(k), \mathbf{X})} & \delta_{\max} < \delta(\mathbf{x}(k), \mathbf{X}). \end{cases} \quad (5.21)$$

Ostane še določitev skupnega veljavnostnega indeksa  $I$ , ki upošteva vseh  $N$  residualov, zbranih v vektorju residualov  $\boldsymbol{\varepsilon}(k)$ . Preprosta možnost je kar srednja

vrednost zadnjih  $N$  veljavnostnih indeksov:

$$I = \frac{I_{k-N+1} + \dots + I_k}{N}, \quad (5.22)$$

seveda pa so na voljo tudi druge.

Izbrana možnost (5.20) za določitev veljavnostnega indeksa ni edina, a je razmeroma preprosta in ponazarja poglobitno idejo.

## Izvedba zaznavanja napak na hidravlični modelni napravi

Zaznavanje napak bomo predstavili na sistemu dveh posod hidravlične modelne naprave [19, 89], predstavljenem v Dodatku E.1. Napravo bomo identificirali z LMGP modelom in predstavili zaznavanje napak na senzorjih izhoda z opisanim algoritmom.

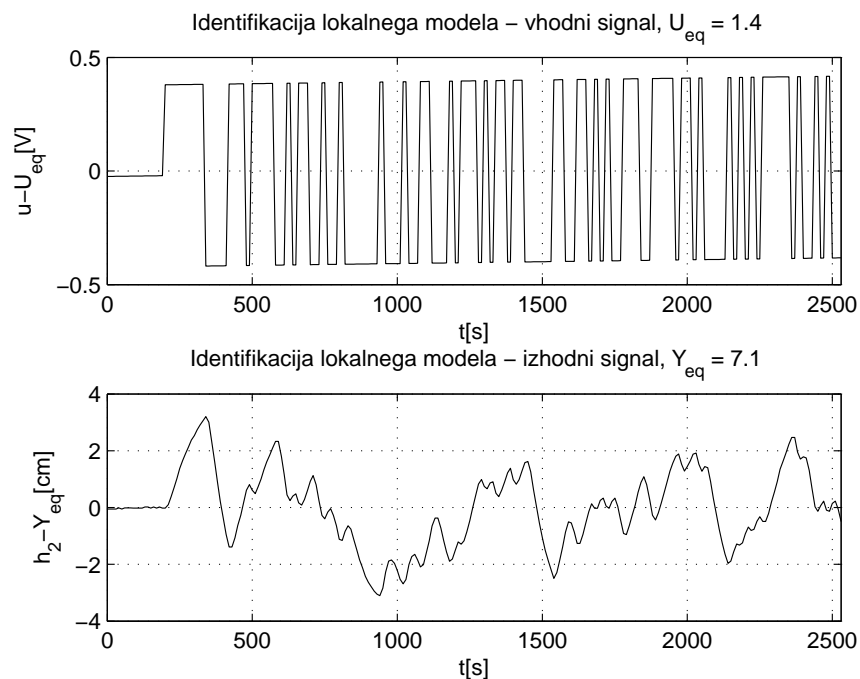
### *Opis področja delovanja naprave in meritev*

Naj ponovimo glavne podatke, ki smo jih že našli v enem izmed prejšnjih podpoglavij. Vhod v Dodatku E.1 opisani sistem drugega reda predstavlja napetost  $U$  [V], ki krmili hitrost  $\omega_1$  črpalke  $P1$ . 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. Statična karakteristika sistema je prikazana na sliki 5.11. Področje delovanja je omejeno z višinama tekočine v posodah R1 in R2, ki ne smeta preseči višine obeh posod. S preizkušanjem smo največjo vrednost napetosti na črpalki P1 omejili na  $U_{\max} = 4V$ .

Uporabljali smo dva tipa meritev:

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

Za pridobitev lokalnega modela v izbrani ravnotežni točki smo sistem s konstantnim vhodnim signalom  $U_{eqi}$  najprej pripeljali v to točko in ga tam vzbujali s frekvenčno dovolj bogatim signalom, da smo iz meritev lahko izluščili dovolj informacije. Z iterativnim postopkom identifikacije pri različnih časih

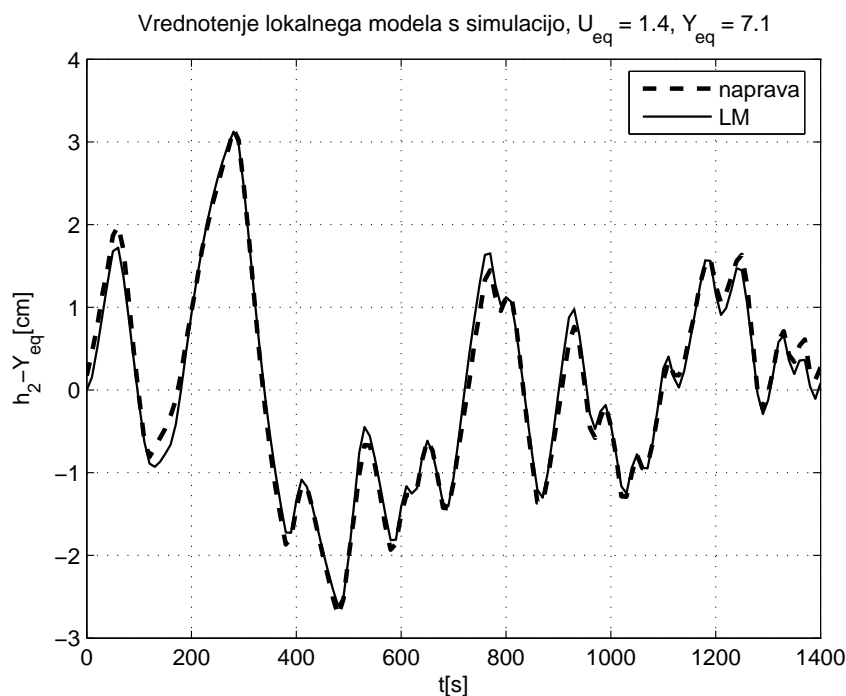


**Slika 5.30:** Signal za identifikacijo lokalnega modela v delovni točki  $(U_{eq}, Y_{eq}) = (1.4, 7.1)$

vzorčenja smo izbrali čas vzorčenja sistema  $T_s = 10$  s. Ko smo sistem pripeljali v delovno točko, smo nosilnemu vhodnemu signalu  $U = U_{eqi}$  dodali pseudo-naključni binarni signal [69] s taktom signala  $T_{sw} = 2T_s = 20$  s in amplitudo  $\Delta U = 0.4$ , ki določa področje veljavnosti modela. Pri izbiri amplitude moramo narediti kompromis: če izberemo manjšo amplitudo, je boljši lokalni opis dinamike, a se hkrati poveča vpliv šuma na identifikacijo. Linearni model je bil v ravnotežni točki identificiran z metodo pomožnih spremenljivk programskega paketa Matlab (`iv4`). Na sliki 5.30 sta prikazana vhodni in izhodni signal za točko  $(U_{eq} = 1.4, Y_{eq} = 7.1)$  in  $\Delta U = 0.4$ .

Enako kot identifikacijski signali so bili dobljeni signali za vrednotenje dobljenih linearnih modelov. Na sliki 5.31 je prikazano vrednotenje linearnega modela s simulacijo v točki  $(U_{eq} = 1.4, Y_{eq} = 7.1)$ .

Drugi tip meritev smo uporabili za pridobivanje učnih podatkov za učenje LMGP modela za področja zunaj ravnotežne krivulje in za preiskus delovanja LMGP modela za zaznavanje napak. Vzbujevalni signali za identifikacijo so bili realizirani podobno kot omenjeni pseudo-naključni binarni signal s taktom oz. najmanjšo spremembo  $T_{sw} = 20$  s, le da je vrednost signala ob spremembi lahko zavzela katerokoli vrednost med  $U_{\min} = 0.8$  in  $U_{\max} = 4$ . Primer vhodnega in



**Slika 5.31:** Signal za vrednotenje lokalnega modela v delovni točki  $(U_{eq}, Y_{eq}) = (1.4, 7.1)$

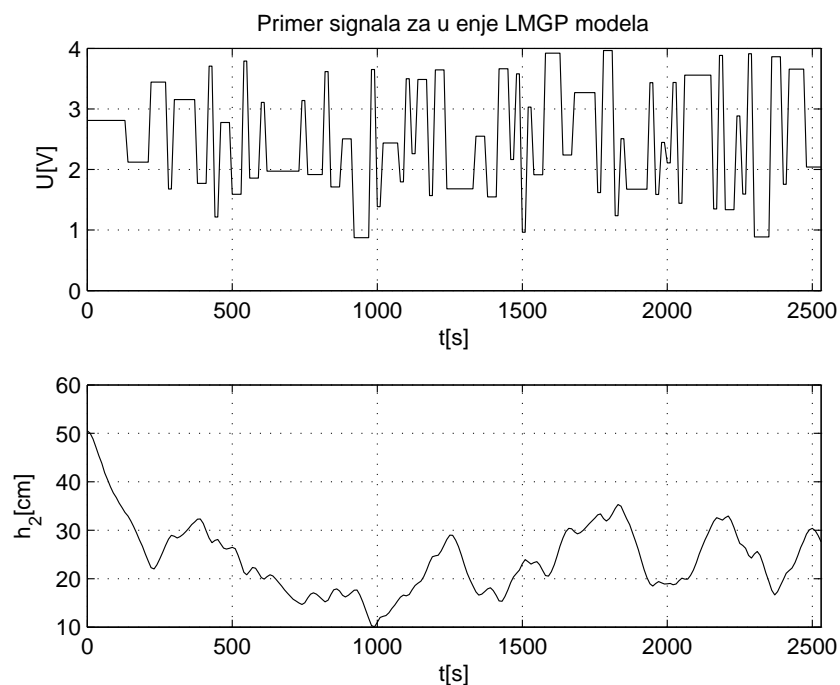
izhodnega signala, iz katerega so bili z vzorčenjem dobljeni podatki za učenje, je prikazan na sliki 5.32.

#### *Identifikacija in vrednotenje LMGP modela za zaznavanje napak*

Radi bi pokazali, kako se LMGP model obnaša pri odkrivanju napak in kako na odločitev o prisotnosti napake vpliva slabša opisanost sistema na določenem področju, zato smo sistem opisali samo na spodnjem delu področja delovanja. Za opis smo izbrali devet linearnih lokalnih modelov, pri čemer je imel lokalni model z največjo vrednostjo izhoda središče v ravnotežni točki ( $U_{eq} = 2.7, Y_{eq} = 26.9$ ). Posamezni lokalni model, ki opisuje obnašanje modela v delovni točki  $(U_{eq}, Y_{eq})$ , opisuje enačba:

$$y(k) = -a_1y(k-1) - a_2y(k-2) + b_1u(k-1) + b_2u(k-2). \quad (5.23)$$

Vsi uporabljeni lokalni modeli in njihovi parametri so zbrani v tabeli 5.8. Diagonalni elementi kovariančnih matrik identificiranih lokalnih modelov, ki predstavljajo varianco identificiranih parametrov, so bili reda velikosti  $10^{-3}$  do  $10^{-4}$ .



**Slika 5.32:** Primer signalov, iz katerih smo dobili podatke za opis delovanja sistema v neravnotežnem področju

Prav tako smo iz učne množice, ki jo uporabimo za opis delovanja sistema v neravnotežnih področjih, odstranili vse učne vektorje, pri katerih je vrednost izhoda modela presegla vrednost  $h_2(k) = 32$  cm. Za opis sistema v neravnotežnem področju smo uporabili 144 učnih točk, če prištejemo temu še devet lokalnih modelov, smo za opis sistema uporabili  $144 + 9 \times 5 = 189$  učnih točk.

**Tabela 5.8:** Parametri uporabljenih lokalnih modelov za model sistema dveh posod

$U_{eq}$	$Y_{eq}$	$-a_1$	$-a_2$	$b_1$	$b_2$
1.2	5.5520	1.3530	-0.4252	0.3821	0.3265
1.4	7.1060	1.3588	-0.4266	0.3592	0.3214
1.6	8.9580	1.4211	-0.4795	0.3799	0.3175
1.8	11.1144	1.4522	-0.5063	0.4456	0.3889
2.0	14.1542	1.5076	-0.5539	0.4484	0.3735
2.2	17.2065	1.5007	-0.5451	0.4525	0.3833
2.3	19.1841	1.5037	-0.5477	0.4597	0.4002
2.5	22.5316	1.5372	-0.5786	0.4730	0.4275
2.7	26.8842	1.5656	-0.6031	0.4368	0.3813

Da dobimo bolj natančno oceno negotovosti napovedi, smo za ovrednotenje modela uporabili simulacijo z numerično aproksimirano propagacijo negotovosti (razdelek 3.5.3, drugi postopek) s  $S = 200$  vzorci. Model smo ovrednotili na dveh področjih:

- na dobro opisanem področju, kar smo dosegli z omejitvijo vhodnega signala  $U$  med  $0.8 < U < 2$  in
- na slabo opisanem področju, omejenim z vrednostmi vhodnega signala  $U$  med  $2.5 < U < 4$ .

Na sliki 5.33 zgoraj je prikazan rezultat simulacije skupaj z napako in 95% pasom zaupanja. Rezultati simulacije na slabo opisanem področju so prikazani na sliki 5.34 in takoj lahko opazimo, da je napovedovanje modela slabše, vendar nas na to opozori ustrezno povečana varianca. Do enakega sklepa pridemo, če opazujemo veljavnostna indeksa  $I$  na obeh slikah. Rezultati obeh simulacij so bili ovrednoteni tudi s cenilkama povprečna kvadratna napaka SE (3.20) in logaritem gostote napake LD (3.22). Za rezultate simulacije na dobro opisanem področju sta vrednosti cenilk  $SE = 0.194$  in  $LD = 1.25$ , za rezultate simulacije na slabo opisanem področju pa  $SE = 4.1$  in  $LD = 2.0$ . Tudi iz teh kvantitavnih ocen lahko sklepamo, da je prvo področje razmeroma dobro opisano, medtem ko je opisnost drugega slabša.

#### *Zaznavanje napak z uporabo LMGP modela*

Tako identificirani in ovrednoteni LMGP model smo uporabili za zaznavanje napak. Pokvarjeni senzor izhoda smo simulirali tako, da smo med simulacijo v času od  $t = 1000$  s do  $t = 1500$  s pravi vrednosti izhoda  $h_2$  prišteli napako v velikosti dveh centimetrov.

V predstavljenem algoritmu za zaznavanje napak moramo glede na sistem nastaviti dva parametra:

1. velikost okna uporabljanih residualov  $M$  in
2. mejno vrednost logaritma razmerja hipotez  $c_{\max}$ .

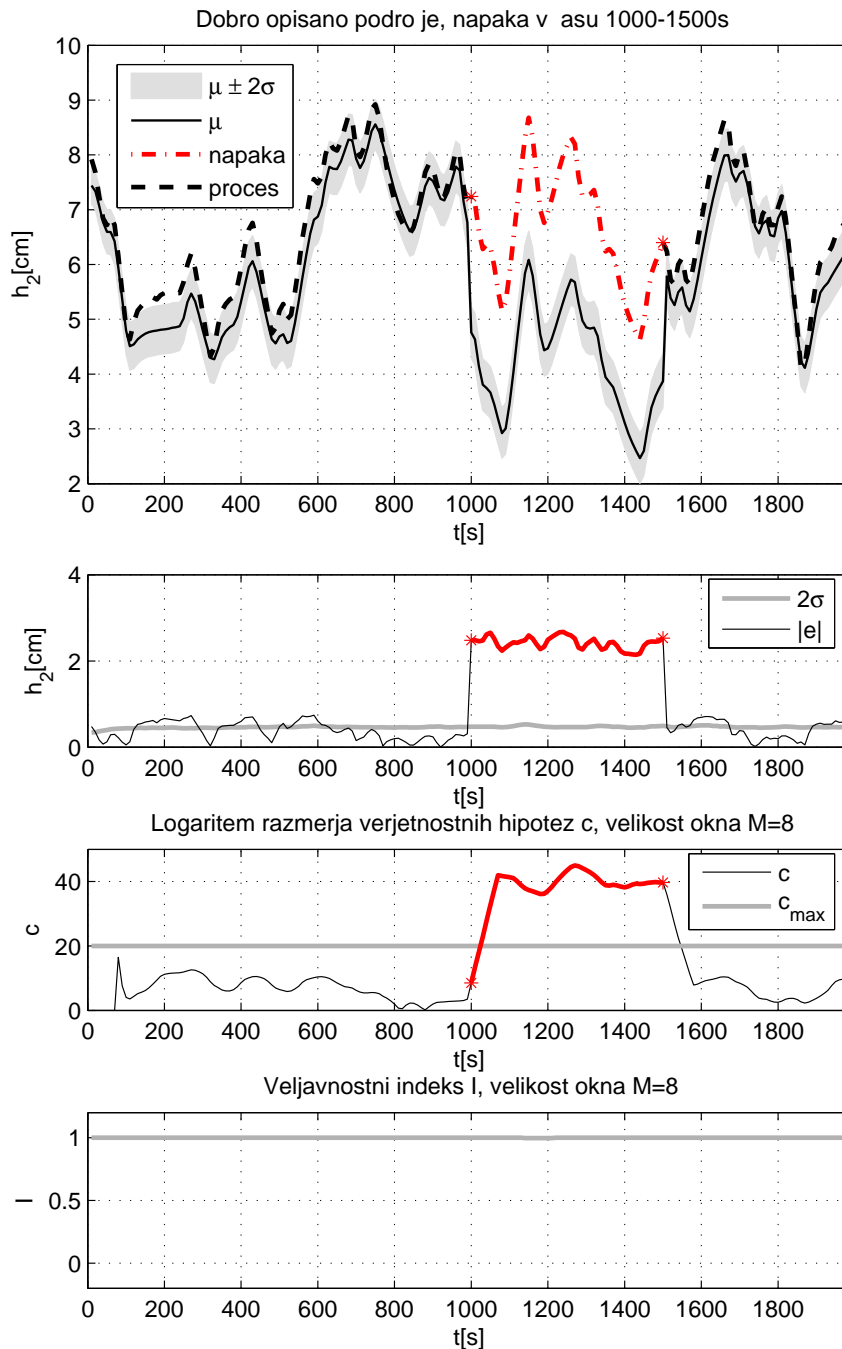
Glede na rezultate vrednotenja smo se odločili za  $M = 8$  in stopnjo pomembnosti  $c = 20$ . Ta je nastavljena razmeroma visoko, saj je modelna naprava izkazovala napako na izhodu senzorjev za nivo med več zaporednimi meritvami. Pri

računanju veljavnostnega indeksa  $I$  smo uporabljali (napovedano) propagirano varianco v  $k$ -tem koraku  $\delta(\mathbf{x}(k), \mathbf{X})$ , pri čemer v skladu s pristopom nismo upoštevali šuma na izhodu.

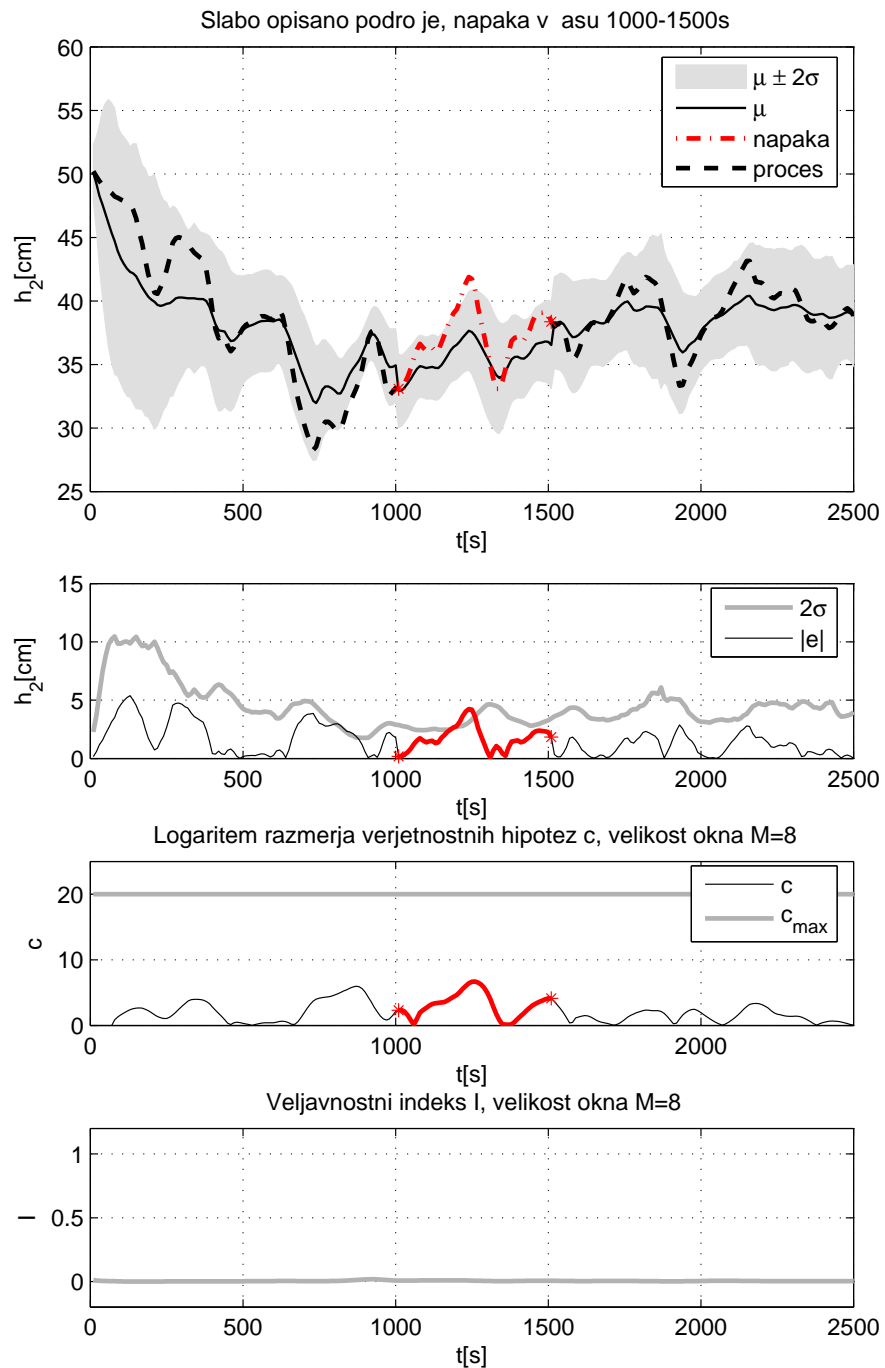
Rezultati algoritma za zaznavanje napak, ko se model nahaja na dobro opisanem področju, so predstavljeni na sliki 5.33. Vidimo, da kritična vrednost logaritma razmerja hipotez preseže mejno vrednost na področju, na katerem smo izhodu sensorja dodali napako, veljavnostni indeks  $I$  pa je vseskozi zelo blizu 1; to pove, da lahko rezultatom algoritma zaupamo. Nasprotno v primeru simulacije modela na slabo opisanem področju (slika 5.34) algoritem napake sensorja ne odkrije, saj napaka zaradi slabšega napovedovanja modela ne pride do izraza. Na slabše napovedovanje lahko sklepamo tudi iz vrednosti veljavnostnega indeksa  $I$ , ki so vseskozi zelo blizu nič (slika 5.34 spodaj), in opozarjajo, da rezultati algoritma za zaznavanje napak nimajo velikega pomena in da je potrebno področje boljše modelirati.

\*       \*       \*

V tem podpoglavju smo pokazali, kako lahko izkoristimo negotovost napovedi GP modela na področju zaznavanja napak. Prvič, natančneje lahko napovemo verjetnostno porazdelitev napake, kar pomaga pri odločanju za najbolj verjetno hipotezo, in drugič, njena uporaba da oceno, koliko lahko dobljenim rezultatom zaupamo, saj varianca napovedi pove, kako dobro je sistem opisan na področju, na katerem je. Te analitične rezultate smo uporabili za uspešno zaznavanje (simulirane) napake izhodnega sensorja hidravlične modelne naprave, identificirane z LMGP modelom.



**Slika 5.33:** Zaznavanje napak z LMGP modelom na dobro opisanem področju, umetno ustvarjena napaka v času od 1000 do 1500 s. Odsek, na katerem je bila simulirana napaka, je označen s pika-črta krivuljo, njen začetek in konec pa dodatno z \*.



**Slika 5.34:** Zaznavanje napak z GP modelom na slabo opisanem področju, umetno ustvarjena napaka v času od 1000 do 1500 s. Odsek, na katerem je bila simulirana napaka, je označen s pika-črta krivuljo, njen začetek in konec pa dodatno z \*. Veljavnostni indeks  $I$  pove, da rezultatom zaznavanja ne moremo zaupati.

## 5.7 Kdaj uporabiti GP model?

V tem poglavju smo predstavili uporabo GP modela in njegovih izpeljank, predstavljenih v četrtem poglavju, za reševanje praktičnih nalog, zaradi katerih modele sploh identificiramo. Dobljeni rezultati kažejo, da so GP modeli privlačni za uporabo v inženirski praksi.

Povzetek do sedaj zbranih napotkov v disertaciji in rezultatov v tem poglavju predstavljenih praktičnih primerov lahko združimo v seznam smernic, ki bodo pomagale pri odločitvi, ali naj za določeno nalogo uporabimo GP model.

Glavna vodila pri odločitvi za uporabo GP modela so:

- sistem, ki bi ga radi opisali, je lahko tudi nelinearen;
- predznanje o sistemu je pomanjkljivo;
- v sistemu ali pri njegovem merjenju je zelo prisoten šum;
- podatki, ki opisujejo sistem, so neenakomerno razporejeni glede na področje delovanja sistema;
- vemo, kako bi lahko koristno uporabili dano zaupanje v napoved izhoda sistema, ki tudi nastopa kot izhod iz modela;
- na voljo je razmeroma majhno (a ne premajhno) število podatkov.

Prvi dve vodili nakazujeta na uporabo identifikacije ne glede na metodo, druga vodila pa usmerjajo k uporabi GP modelov.

Tipični primeri sistemov, o katerih imamo manj predznanja in/ali je njihovo modeliranje oteženo zaradi nelinearnosti, neznanih ali težko merljivih vplivov, so biotehnološki, biološki, kemijski, meteorološki sistemi itd.

Zaupanje v napoved je ena izmed lastnosti, ki GP model najbolj loči od večine drugih načinov identifikacije z modelom črne škatle. Pove, koliko lahko napovedi modela zaupamo glede na identifikacijske podatke, kar je dobrodošla dodatna informacija pri napovedovanju obnašanja sistemov. Koristno bi jo lahko uporabili tudi pri načrtovanju vodenja, npr. za načrtovanje robustnega vodenja, izračun optimalne vrednosti regulirne veličine (minimalno variančno vodenje) itd. Na področju zaznavanja napak lahko zaupanje v napoved uporabimo kot informacijo o kvaliteti opisa področja delovanja sistema, ki je lahko npr. pomoč pri odločanju o sproženju alarma, s čimer povečamo robustnost zaznavanja. Informacija o kvaliteti opisa določenega področja delovanja je koristna tudi pri samem

vrednotenju modela, ker opozori na slabše opisana področja. Na njeni podlagi se lahko odločimo za zbiranje dodatne informacije o sistemu in ponovnem učenju modela ali pa, če to ni mogoče, model zavržemo.

Če imamo na voljo predznanje o sistemu, tega lahko vključimo v GP model, kot je bilo pokazano v četrtem poglavju. Tam smo posebej poudarili dve izpeljanki GP modela: GP model z vključenimi lokalnimi modeli (LMGP model) in GP model z vnaprej določeno strukturo (FSGP model).

LMGP model omogoča možnost enostavnega združevanja informacije o lokalnem obnašanju sistema v obliki lokalnih modelov. Še posebej pomembno je, če želimo že znani informaciji o sistemu dodati novo, naj bo to informacija v obliki lokalnih modelov ali v obliki vzorčenih podatkov, iz katerih je sicer sestavljen GP model, saj je združevanje informacije relativno enostavno. Dodatna ugodnost LMGP modela je zaradi uporabe lokalnih modelov namesto vzorčenih podatkov navadno manjša učna množica, kar potencialno zmanjša računsko zahtevnost in pohitri optimizacijo hiperparametrov ter napovedovanje takega modela. Tudi LMGP model je neparametričen, uporabimo ga:

- kadar bi radi v model, sestavljen iz linearnih lokalnih modelov, ki ponavadi opisujejo dinamiko blizu ravnotežnih področij, vstavili (pomanjkljive) podatke s področja zunaj ravnotežja;
- kadar želimo zmanjšati količino podatkov v ravnotežnih točkah;
- kadar je za opis neravnotežnega področja zadostno manjše število točk.

Tudi v LMGP modelu lahko spremljamo razširjanje negotovosti predikcije pri simulaciji.

Posebna oblika uporabe GP modela je GP model z vnaprej določeno strukturo (FSGP model). Kar ta model najbolj razlikuje od drugih "izpeljank" GP modelov je njegova *parametričnost*, ki je koristna, če hočemo model uporabiti npr. za načrtovanje vodenja. Problem FSGP modela, ki je skupen večini pristopov (izjema je npr. LMGP model), ki modelirajo globalno obnašanje sistema s primernim združevanjem opisa lokalnega obnašanja, se pojavi, če je potrebno opisati sistem tudi na področjih, odmaknjenih od ravnotežne krivulje. V tem primeru smo: primorani identificirati lokalne modele tudi na teh področjih, kar je navadno znatno težje, včasih tudi nemogoče, ali pa predpostavljati, da so ta področja zadosti dobro opisana le z lokalnimi modeli blizu ravnotežne krivulje, kar vodi do problemov, opisanih v podpoglavju 3.1. Hitrost optimizacije GP modelov parametrov, ki sestavljajo FSGP model, ni problematična, saj nanjo mnogo bolj kot red modela vpliva število učnih točk. Pri večjih redih je bolj problematično, da moramo za dober FSGP model identificirati razmeroma veliko število lokalnih

modelov, ki jih potrebujemo za učenje. Glavno vodilo pri izbiri je potreba po parametričnem modelu, dodatna ugodnost FSGP modela pa je, da njegov izhod poleg napovedane vrednosti izhoda sistema vključuje tudi mero negotovosti za napovedane vrednosti parametrov predpostavljenega modela z vnaprej določeno strukturo.

Našteti so bili razlogi za uporabo GP modela, poglejmo še, kdaj njihova uporaba ni smiselna. V primeru, da dinamični sistem za naš namen lahko zadovoljivo opišemo z linearnim modelom ali z modelom zaradi predznanja poznane strukture, ki mu samo optimiramo neznane parametre, GP modela za opis sistema verjetno ne potrebujemo. Prav tako "običajnega" GP modela ne bomo uporabili, če potrebujemo parametrični model sistema ali če bi želeli model, ki odraža fizikalno ozadje. Neprijetnost, ki se zazda pojavlja pri uporabi GP modela, je razmeroma velika računsko zahtevnost tako pri učenju modela kot pri njegovi simulaciji, posebno pri velikih modelih, tj. pri modelih z veliko učno množico. Sicer obstajajo metode, ki lahko pospešijo delovanje GP modela (omenili smo jih na strani 27), prav tako lahko predvidevamo, da se bo s tehničnim napredkom računsko moč, ki jo imamo na voljo, povečala. Vseeno je trenutno treba pri aplikacijah, ki zahtevajo veliko računsko hitrost, GP modele uporabljati premišljeno. Zaradi narave in hitrostnih omejitev GP modela je problematična tudi uporaba GP modela kot adaptivnega, saj dodatna informacija, ki jo dobivamo zahteva ponovno učenje modela in hkrati povzroči povečevanje obsega modela.

Naj na koncu poudarimo, da je kvaliteta GP modela tako kot kvaliteta drugih modelov odvisna od kvalitete in obsega uporabljene učne informacije. V splošnem ne moremo pričakovati, da bo dal GP model boljše predikcije od drugih modelov, dobimo pa informacijo o zaupanju, ki jo lahko primerno izkoristimo.

## Poglavje 6

# Zaključek

V disertaciji smo obravnavali identifikacijo nelinearnih dinamičnih sistemov z modeli na osnovi Gaussovih procesov (GP modeli).

GP model je verjetnostni, neparametrični model, ki napoveduje izhod v obliki Gaussove porazdelitve. Po predstavitvi GP modela in njegove uporabe za reševanje regresijskih problemov smo na tej podlagi postavili praktični okvir za identifikacijo dinamičnih sistemov in podrobneje obravnavali posamezne faze postopka. Poseben poudarek smo namenili:

- postavitvi GP modela, pri čemer smo dali napotke za izbiro kovariančne funkcije, za izbiro regresorjev in posredno reda modeliranega sistema. Pri izbiri kovariančne funkcije upoštevamo predznanje o sistemu, pri izbiri regresorjev pa si poleg predznanja lahko pomagamo še z vrednotenjem modela in lastnostjo GP modela, da se pomembnost posameznih regresorjev odraža skozi vrednost ustreznih hiperparametrov;
- vrednotenju GP modela, pri tem smo poleg bolj uporabljenih načinov vrednotenja pokazali tudi, kako pri vrednotenju koristno uporabimo za GP model značilno varianco napovedanega izhoda, in
- simulaciji GP modela: pokazali smo, kako simulirati GP model s propagacijo negotovosti izhoda ali brez nje. Propagacija negotovosti da natančnejši rezultat, a je težavnejša za implementacijo in računsko zahtevnejša, zato jo uporabimo samo tam, kjer nas zanima večja natančnost. Če zadostuje že informacija o kvaliteti opisa področja, na katerem je sistem (to informacijo nosi povečanje ali zmanjšanje vrednosti variance), potem uporabimo raje hitrejšo in enostavnejšo simulacijo brez propagacije negotovosti izhoda.

Drugi večji sklop disertacije je bilo vključevanje morebitnega predznanja o identificiranem sistemu v GP model. Predznanje lahko vključujemo na dva različna načina: s spreminjanjem kovariančne funkcije oz. matrike in s spreminjanjem vhodno/izhodnih podatkov, ki sestavljajo GP model. Povzeli smo vključevanje:

- statične karakteristike in
- lokalnih modelov

ter pokazali kako vključevati:

- znanje o šumnem procesu kot motnji na izhodu identificiranega sistema,
- histerezo kot specifičnega predstavnika nelinearnosti, ki se pogosto pojavlja v praksi in
- znanje o strukturi opisovanega sistema, ki je botrovalo razvoju GP modela z vnaprej določeno strukturo.

Ugotovili smo, da ima vključevanje predznanja pozitiven vpliv tako na kvaliteto modela kot na postopek identifikacije. Vključevanje predznanja izboljša napovedovanje modela, vključevanje lokalnih modelov pa lahko tudi močno zmanjša učno množico, kar pripomore k hitrejšemu učenju in napovedovanju modela. Iz prikazanih primerov lahko izluščimo tudi, kako bi postopali, če bi vključevali druge vrste predznanja, ki jih v disertaciji nismo posebej obravnavali.

V zadnjem delu disertacije smo na praktičnih primerih ilustrirali v začetnih poglavjih predstavljeno problematiko disertacije. Prikazali smo:

- postopek identifikacije GP modela s poudarkom na njegovem vrednotenju,
- kako poteka vključevanje lokalnih modelov v GP model in kako uporabiti znanje o strukturi opisovanega sistema za identifikacijo GP modela z vnaprej določeno strukturo in
- ilustrirali, kako lahko GP modele uporabimo za filtriranje, napovedovanje odziva, načrtovanje vodenja in zaznavanje napak.

Navedli smo tudi glavna vodila, na podlagi katerih se odločimo za uporabo GP modela: sistem, ki bi ga radi opisali, je lahko tudi nelinearen, predznanje o sistemu je pomanjkljivo, v sistemu ali pri njegovem merjenju je zelo prisoten šum, podatki so neenakomerno razporejeni glede na delovanje sistema, lahko jih

je tudi razmeroma malo. Prav tako je smiselna uporaba GP modela, če je mogoče dano zaupanje v napoved izhoda sistema koristno uporabiti.

Če imamo na voljo predznanje o sistemu, tega seveda lahko vključimo v GP model. Posebna primera vključevanja predznanja sta GP model z vključenimi lokalnimi modeli in GP model z vnaprej določeno strukturo. Prvi je pomemben, ker ponuja možnost enostavnega združevanja informacije v obliki lokalnih modelov in vzorčenih podatkov, drugi pa zaradi parametričnosti, kar lahko koristno uporabimo npr. pri načrtovanju vodenja, pri čemer kot dodatno ugodnost poleg napovedanega izhoda sistema dobimo še mero negotovosti za vrednosti napovedanih parametrov predpostavljenega modela vnaprej določene strukture.

Med delom z GP modeli se je pokazalo, da je predvsem učenje modela z obsežno učno množico in velikim številom regresorjev časovno potratno. Ena izmed možnih rešitev je zmanjšanje učne množice z omenjenim združevanjem informacije v lokalne modele in z njihovim vključevanjem v model, druga (podobna) pa je uporaba naprednejših algoritmov, ki bi pohitrili delovanje modela. Med omejitve GP modela spada njegova neparametričnost in potreba po podatkih, ki opisujejo celotno področje delovanja sistema. Te omejitve niso tako hude, saj je zlasti v inženirski praksi pogosta uporaba neparametričnih modelov, na slabo opisana področja sistema pa uporabnika opozori varianca napovedanega izhoda GP modela.

Predstavljeni primeri so pokazali, da bi GP modele lahko uporabljali v praktičnih aplikacijah tudi v industrijskem okolju, kar se bo, upamo, sčasoma tudi uveljavilo. GP modeli kažejo svojo potencialno uporabnost v sistemih za zaznavanje napak in pri načrtovanju robustnega vodenja.

Dodatno delo na področju identifikacije dinamičnih sistemov z GP modeli bi bilo potrebno nameniti:

- razvoju uporabniku prijaznejše in popolnejše programske opreme za pomoč pri identifikaciji in simulaciji,
- določanju začetnih vrednosti hiperparametrov kovariančne funkcije pri optimizaciji,
- povečanju računske učinkovitosti pri identifikaciji,
- razvoju orodja za analizo lastnosti GP modela v zaprtozančnih sistemih (npr. za analizo stabilnosti);
- potrebno je nadaljevati raziskave o uporabi GP modelov v vodenju, pri čemer bi lahko koristno uporabili mero negotovosti napovedi izhoda sistema, in

- ter raziskati in potencialno uporabiti modeliranje z GP modeli za opis sistemov, ki imajo v svojem obnašanju inherentno nedoločenost (npr. modeliranje človeškega obnašanja).
- Poleg naštetega dela, ki velja za identifikacijo GP modelov v splošnem, bi bilo potrebno raziskati, ali lahko izboljšamo postopek učenja GP modelov, ki sestavljajo GP model z vnaprej določeno strukturo, z uporabo informacije o potencialni medsebojni odvisnosti parametrov lokalnih modelov, ki jih GP modeli opisujejo.

Raziskovanje je dejavnost, ki z vsakim odgovorom na vprašanje navadno odpre nova vprašanja. Odkrivanje zakonitosti in možnosti uporabe modelov na osnovi Gaussovih procesov za modeliranje dinamičnih sistemov ni nobena izjema, a vsako odgovorjeno vprašanje ali razkrita zakonitost sta lahko nov korak v prenosu dognanj v inženirsko prakso.

## Literatura

- [1] F. Allgöwer, Editorial: Nonlinear model predictive control, *IEE Proceedings Control Theory and Applications*, Vol. 152, No. 3, str. 257–258, 2005.
- [2] F. Allgöwer, Z. Nagy, R. Findeisen, Nonlinear model predictive control: From theory to application, *Proceedings of International Symposium on Design, Operation and Control of Chemical Plants, (PSE)*, str. 639–650, Taipei, Taiwan, 2002.
- [3] A. Aoyama, F. J. Doyle, V. Venkatasubramanian, A fuzzy neural networks approach for nonlinear process control, *Engineering Applications of Artificial Intelligence*, Vol. 8, No. 5, str. 483–498, 1995.
- [4] K. Ažman, *Identifikacija dinamičnih sistemov z Gaussovimi procesi z vključenimi lokalnimi modeli*, magistrsko delo, Univerza v Ljubljani, Ljubljana, september 2004.
- [5] K. Ažman, Incorporating prior knowledge into Gaussian process model, *Proceedings of 6th International PhD Workshop on Systems and Control – A Young Generation Viewpoint*, Volume A, str. 253–256, Izola, 2005.
- [6] K. Ažman, J. Kocijan, Comprising prior knowledge in dynamic Gaussian process models, *Proceedings of the International Conference on Computer Systems and Technologies – CompSysTech, IIB.2*, Varna, 2005.
- [7] K. Ažman, J. Kocijan, An example of Gaussian process model identification, L. Budin, S. Ribarić, uredniki, *Proceedings of 28th International conference MIPRO, CIS – Intelligent Systems*, str. 79–84, Opatija, maj 2005.
- [8] K. Ažman, J. Kocijan, Identifikacija dinamičnega sistema s histerezo z modelom na osnovi Gaussovih procesov, B. Zajc, A. Trost, uredniki, *Zbornik štirinajste elektrotehniške in računalniške konference ERK*, Volume A, str. 253–256, Portorož, 2005.

- [9] K. Ažman, J. Kocijan, An application of Gaussian process models for control design, *UKACC International Control Conference*, str. 1092–1123, Glasgow, 2006.
- [10] K. Ažman, J. Kocijan, Gaussian process model validation: biotechnological case studies, I. Troch, F. Breitenecker, uredniki, *Proceedings of the 5th Vienna Symposium on Mathematical Modelling – MathMod*, Dunaj, 2006.
- [11] K. Ažman, J. Kocijan, Identifikacija dinamičnega sistema z znanim modelom šuma z modelom na osnovi Gaussovih procesov, B. Zajc, A. Trost, uredniki, *Zbornik petnajste elektrotehniške in računalniške konference ERK*, Volume A, str. 289–292, Portorož, 2006.
- [12] K. Ažman, J. Kocijan, Application of Gaussian processes for black-box modelling of biosystems, *ISA Transactions*, Vol. 46, No. 4, str. 443–457, oktober 2007, <http://dx.doi.org/10.1016/j.isatra.2007.04.001>.
- [13] K. Ažman, J. Kocijan, Nonlinear model predictive control for models with local information and uncertainties, *The Transactions of the Institute of Measurement and Control*, 2007, sprejeto v objavo.
- [14] B. Bamieh, L. Giarré, Identification of linear parameter varying models, *International Journal Of Robust And Nonlinear Control*, Vol. 12, No. 9, str. 841–853, 2002.
- [15] C. M. Bishop, M. E. Tipping, Bayesian regression and classification, J. Suykens, G. Horvath, S. Basu, C. Micchelli, J. Vandewalle, uredniki, *Advances in Learning Theory: Methods, Models and Applications*, Volume 190 iz *NATO Science Series III: Computer and Systems Sciences*, Amsterdam, 2003, IOS Press.
- [16] J. Cho, J. C. Principe, D. Erdogmus, M. A. Motter, Quasi-sliding model control strategy based on multiple-linear models, *Neurocomputing*, Vol. 70, str. 960–974, 2007.
- [17] G. Coffaro, G. Carrer, G. Bendoricchio, Model for *Ulva rigida* growth in the lagoon of Venice, Delovno poročilo, UNESCO MURST Project: Venice Lagoon Ecosystem, University of Padua, Padova, Italija, 1993.
- [18] *Columbia Electronic Encyclopedia*, Columbia University Press, šesta izdaja, 2003, <http://www.answers.com/topic/hysteresis>, pridobljeno 27.08.2007 s svetovnega spleta.
- [19] G. Dolanc, Dj. Juričić, A. Rakar, J. Petrovčič, D. Vrančič, Three-tank benchmark test, Delovno poročilo CT94-02337, Copernicus Project, Institut Jožef Stefan, Ljubljana, 1997.

- 
- [20] R. Dybowski, V. Gant, *Clinical Applications of Artificial Neural Networks*, Cambridge University Press, Cambridge, 2001.
- [21] ELWE, ELWE educational systems, Experimental manual U9, ELWE, Cremlingen 1988.
- [22] A. Emami-Naeini, M. M. Akhter, S. M. Rock, Effect of model uncertainty on failure detection: the threshold selector, *IEEE Transactions on Automatic Control*, Vol. 33, No. 12, str. 1106–1115, 1988.
- [23] L. Fausett, *Fundamentals of neural networks: architectures, algorithms, and applications*, Prentice-Hall, New Jersey, 1994.
- [24] M. W. Gardner, S. R. Dorling, Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences, *Atmospheric Environment*, Vol. 32, No. 14-15, str. 2627–2636, 1998.
- [25] J. J. Gertler, *Fault detection and diagnosis in engineering systems*, Marcel Dekker, New York, 1998.
- [26] M. N. Gibbs, *Bayesian Gaussian processes for regression and classification*, doktorska disertacija, Cambridge University, Cambridge, 1997.
- [27] A. Girard, *Approximate methods for propagation of uncertainty with Gaussian process models*, doktorska disertacija, University of Glasgow, Glasgow, 2004.
- [28] A. Girard, C. E. Rasmussen, R. Murray-Smith, Gaussian process priors with uncertain inputs: multiple-step-ahead prediction, Delovno poročilo DCS TR-2002-119, University of Glasgow, Glasgow, 2002.
- [29] P. W. Goldberg, C. K. I. Williams, C. M. Bishop, Regression with input-dependent noise: A Gaussian process treatment, M. I. Jordan, M. J. Kearns, S. A. Solla, uredniki, *Advances in Neural Information Processing Systems*, Volume 10, MIT Press, MA, 1998.
- [30] C. G. Goodwin, S. F. Graebe, M. E. Salgado, *Control system design*, Prentice Hall, New Jersey, 2001.
- [31] B. Grašič, *Napovedovanje povišanih koncentracij ozona z uporabo umetnih nevronskih mrež, Gaussovih procesov in mehke logike*, magistrsko delo, Univerza v Ljubljani, Ljubljana, 2005.
- [32] G. Gregorčič, *Data-based modelling of nonlinear systems for control*, doktorska disertacija, University College Cork, National University of Ireland, Cork, 2004.

- [33] G. Gregorčič, G. Lightbody, Gaussian processes for modelling of dynamic non-linear systems, *Proceedings of Irish Signals and Systems Conference, Cork*, str. 141–147, Cork, junij 2002.
- [34] G. Gregorčič, G. Lightbody, An affine Gaussian process approach for nonlinear system identification, *Systems Science Journal*, Vol. 29, No. 2, str. 47–63, 2003.
- [35] G. Gregorčič, G. Lightbody, From multiple model networks to the Gaussian processes prior model, *Proceedings of IFAC ICONS conference*, str. 149–154, Faro, 2003.
- [36] N. Hvala, S. Strmčnik, D. Šel, S. Milanič, B. Banko, Influence of model validation on proper selection of process models — an industrial case study, *Computers and Chemical Engineering*, Vol. 29, str. 1507–1522, 2005.
- [37] T. A. Johansen, R. Shorten, R. Murray-Smith, On the interpretation and identification of dynamic Takagi-Sugeno fuzzy models, *IEEE Transactions on Fuzzy Systems*, Vol. 8, No. 3, str. 297–313, 2000.
- [38] K. Judd, M. Small, Towards long-term prediction, *Physica D*, Vol. 136, No. 1-2, str. 31–44, 2000.
- [39] A. Juditsky, H. Hjalmarsson, A. Benveniste, B. Delyon, L. Ljung, J. Sjöberg, Q. Zhang, Nonlinear black-box models in system identification: Mathematical foundations, *Automatica*, Vol. 31, No. 12, str. 1725–1750, 1995.
- [40] Dj. Juričić, J. Kocijan, Fault detection based on Gaussian process model, I. Troch, F. Breitenecker, uredniki, *Proceedings of the 5th Vienna Symposium on Mathematical Modelling – MathMod*, Dunaj, 2006.
- [41] M. Kinnaert, Fault diagnosis based on analytical models for linear and nonlinear systems: a tutorial, *Preprints of IFAC symposium SAFEPROCESS*, str. 17–50, Washington, 2003.
- [42] G. Klančar, Dj. Juričić, R. Karba, Robust fault detection based on compensation of the modelling error, *International Journal of System Science*, Vol. 33, No. 2, str. 97–105, 2002.
- [43] J. Kocijan, *Praktikum identifikacij dinamičnih sistemov*, Fakulteta za elektrotehniko, Ljubljana, 1998.
- [44] J. Kocijan, K. Ažman, Gaussian process model identification: A process engineering case study, A. Grzech, urednik, *Proceedings of 16th International Conference on Systems Science*, Volume 1, Wroclaw, september 2007.

- 
- [45] J. Kocijan, B. Banko, B. Likar, A. Girard, R. Murray-Smith, C. E. Rasmussen, A case based comparison of identification with neural networks and Gaussian process models, *Proceedings of IFAC ICONS conference*, str. 137–142, Faro, 2003.
- [46] J. Kocijan, A. Girard, Incorporating linear local models in Gaussian process model, *Proceedings of IFAC 16th World Congress*, Praga, 2005.
- [47] J. Kocijan, A. Girard, B. Banko, R. Murray-Smith, Dynamic systems identification with Gaussian processes, *Mathematical and Computer Modelling of Dynamic Systems*, Vol. 11, No. 4, str. 411–424, december 2005.
- [48] J. Kocijan, A. Girard, D. J. Leith, Incorporating linear local models in Gaussian process model, Delovno poročilo DP-8895, Institut Jožef Stefan, Ljubljana, december 2003.
- [49] J. Kocijan, D. J. Leith, Derivative observations used in predictive control, *Proceedings of Melecon 2004*, Volume 1, str. 379–382, Dubrovnik, 12.-15. maj 2004.
- [50] J. Kocijan, R. Murray-Smith, C. E. Rasmussen, A. Girard, Gaussian process model based predictive control, *Proceedings of 4th American Control Conference*, str. 2214–2218, Boston, MA, 30. junij-2. julij 2004.
- [51] J. Kocijan, G. Žunič, S. Strmčnik, D. Vrančič, Fuzzy gain-scheduling control of gas-liquid separation plant implemented on PLC, *International Journal of Control*, Vol. 75, No. 14, str. 1082–1091, 2002.
- [52] D. G. Krige, A statistical approach to some basic mine valuation problems on the witwatersrand, *Journal of Chemistry, Metal. and Mining Soc. of South Africa*, Vol. 52, No. 6, str. 119–139, 1951.
- [53] D. J. Leith, W. E. Leithead, Analytic framework for blended multiple model systems using linear local models, *International Journal of Control*, Vol. 72, str. 605–619, 1999.
- [54] D. J. Leith, W. E. Leithead, Gain-scheduled control: relaxing slow variation requirements by velocity-based design, *Journal of Guidance, Control and Dynamics*, Vol. 23, No. 6, str. 988–1000, 2000.
- [55] D. J. Leith, W. E. Leithead, Global reconstruction of nonlinear systems from families of linear systems, *Proceedings of World IFAC Congress*, Barcelona, 2002.
- [56] D. J. Leith, W. E. Leithead, E. Solak, R. Murray-Smith, Divide and conquer identification using Gaussian processes, *Proceedings of the 41st Conference on Decision and Control*, str. 624–629, Las Vegas, AZ, 2002.

- [57] D. J. Leith, A. Tsourdos, B. A. White, W. E. Leithead, Application of velocity-based gain-scheduling to lateral auto-pilot design for an agile missile, *Control Engineering Practice*, Vol. 9, str. 1079–1093, 2001.
- [58] W. E. Leithead, K. S. Neo, D. J. Leith, Gaussian regression based on models with two stochastic processes, *Proceedings of IFAC 16th World Congress*, Praga, 2005.
- [59] J. A. Leonard, M. A. Kramer, L. H. Ungar, Using radial basis functions to approximate a function and its error bounds, *IEEE Transactions on Neural Networks*, Vol. 3, No. 4, str. 624–627, julij 1992.
- [60] G. Lightbody, G. W. Irwin, Nonlinear control structures based on embedded neural system models, *IEEE Transactions on Neural Networks*, Vol. 8, str. 533–567, maj 1997.
- [61] B. Likar, *Prediktivno vodenje nelinearnih sistemov na osnovi Gaussovih procesov*, magistrsko delo, Univerza v Ljubljani, Ljubljana, september 2004.
- [62] B. Likar, J. Kocijan, Predictive control of a gas-liquid separation plant based on a Gaussian process model, *Computers and Chemical Engineering*, Vol. 31, str. 142–152, 2007.
- [63] I. Lind, *Regressor selection in system identification using ANOVA*, doktorska disertacija, University of Linköping, Linköping, 2001.
- [64] L. Ljung, *System identification – theory for the user*, Prentice Hall, New Jersey, druga izdaja, 1999.
- [65] D. J. C. MacKay, Introduction to Gaussian processes, C. M. Bishop, urednik, *Neural networks and machine learning*, NATO ASI Series – F 168, str. 133–166, Berlin, 1998, Springer-Verlag.
- [66] D. J. C. MacKay, *Information theory, inference and learning algorithms*, poglavje Gaussian Processes, str. 535–548, Cambridge University Press, Cambridge, 2003.
- [67] E. H. Mamdani, Application of fuzzy logic to approximate reasoning using linguistic systems, *Fuzzy Sets and Systems*, Vol. 26, str. 1182–1191, 1977.
- [68] G. Marsaglia, W. W. Tsang, The ziggurat method for generating random variables, <http://www.jstatsoft.org/v05/i08/ziggurat.pdf>, pridobljeno 19.12.2006 s svetovnega spleta.
- [69] D. Matko, *Identifikacije*, Fakulteta za elektrotehniko, Ljubljana, druga izdaja, 1998.

- 
- [70] D. Matko, G. Geiger, T. Werner, Neural net versus classical models for the detection and localization of leaks in pipelines, *Mathematical and Computer Modelling of Dynamical Systems*, Vol. 12, No. 6, str. 505–517, december 2006.
- [71] D. Matko, I. Škrjanc, G. Mušič, Robustness of fuzzy control and its application to a thermal plant, *Mathematics and Computer in Simulation*, Vol. 51, str. 245–255, 2000.
- [72] D. J. Murray-Smith, Methods for the external validation of continuous system simulation models: A review, *Mathematical and Computer Modelling of Dynamical Systems*, Vol. 4, str. 5–31, 1998.
- [73] R. Murray-Smith, A. Girard, Gaussian process priors with ARMA noise models, *Proceedings of Irish Signals and Systems Conference*, str. 147–152, Maynooth, 2001.
- [74] R. Murray-Smith, T. A. Johansen, uredniki, *Multiple model approaches to modelling and control*, Taylor and Francis, London, 1997.
- [75] R. Murray-Smith, T. A. Johansen, R. Shorten, On transient dynamics, off-equilibrium behaviour and identification in blended multiple model structures, *Proceedings of European Control Conference*, BA-14, Karlsruhe, 1999.
- [76] R. Murray-Smith, D. Sbarbaro, Nonlinear adaptive control using nonparametric gaussian process prior models, *Proceedings of IFAC 15th World Congress*, Barcelona, 2002.
- [77] R. Murray-Smith, D. Sbarbaro, C. E. Rasmussen, A. Girard, Adaptive, cautious, predictive control with Gaussian process priors, *Proceedings of 13th IFAC Symposium on System Identification*, str. 1195–1200, Rotterdam, 2003.
- [78] K. S. Narendra, K. Parthasarathy, Identification and control of dynamical systems using neural networks, *IEEE Transactions on Neural Networks*, Vol. 1, No. 1, str. 4–27, 1990.
- [79] R. M. Neal, *Bayesian learning for neural networks*, Volume 118 iz *Lecture Notes in Statistics*, Springer-Verlag, New York, 1996.
- [80] R. M. Neal, Monte carlo implementation of Gaussian process models for Bayesian regression and classification, Delovno poročilo 9702, Department of Statistics, University of Toronto, 1997, <http://www.cs.toronto.edu/~radford>.

- [81] A. O'Hagan, On curve fitting and optimal design for regression (with discussion), *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 40, No. 1, str. 1–42, 1978.
- [82] G. Olsson, B. Newell, *Wastewater treatment systems. Modelling, diagnosis and control*, IWA Publishing, London, 1999.
- [83] A. Papoulis, *Probability, random variables and stochastic processes*, McGraw-Hill, New York, tretja izdaja, 1991.
- [84] N. Pavešić, *Razpoznavanje vzorcev*, Fakulteta za elektrotehniko, Ljubljana, 1992.
- [85] R. K. Pearson, Selecting nonlinear model structures for computer control, *Journal of Process Control*, Vol. 13, No. 1, str. 1–26, februar 2003.
- [86] J. C. Principe, N. R. Euliano, W. C. Lefebvre, *Neural and adaptive systems – fundamentals through simulations*, John Wiley and Sons, New York, 2000.
- [87] J. Quiñonero-Candela, *Learning with uncertainty – Gaussian processes and relevance vector machines*, doktorska disertacija, Technical University of Denmark, Kongens Lyngby, 2004.
- [88] J. Quiñonero-Candela, C. E. Rasmussen, *Switching and Learning in Feedback Systems*, Volume 3355 iz *Lecture Notes in Computer Science*, poglavje Analysis of Some Methods for Reduced Rank Gaussian Process Regression, str. 98–127, Springer, Heidelberg, 2005.
- [89] A. Rakar, *Odkrivanje napak v tehničnih sistemih z metodami aproksimativnega sklepanja*, doktorska disertacija, Fakulteta za elektrotehniko, Univerza v Ljubljani, Ljubljana, 2000.
- [90] C. E. Rasmussen, *Evaluation of Gaussian processes and other methods for nonlinear regression*, doktorska disertacija, University of Toronto, Toronto, 1996.
- [91] C. E. Rasmussen, C. K. I. Williams, *Gaussian Processes for machine learning*, The MIT Press, Cambridge, MA, 2006.
- [92] V. K. Rohatgi, *An introduction to probability theory and mathematical statistics*, John Wiley and Sons, New York, 1976.
- [93] Ye. N. Rosenvasser, E. Yu. Polyakov, B. P. Lampe, Application of Laplace transformation for digital redesign of continuous control systems, *IEEE Transactions on Automatic Control*, Vol. 44, No. 4, str. 883–886, april 1999.
- [94] W. Rugh, J. Shamma, Research on gain scheduling, *Automatica*, Vol. 36, str. 1401–1425, 2000.

- [95] D. Sbarbaro, R. Murray-Smith, A. Valdes, Multivariable generalized minimum variance control based on artificial neural networks and gaussian process models, *International Symposium on Neural Networks*, Springer Verlag, 2004.
- [96] M. Seeger, C. K. I. Williams, N. D. Lawrence, Fast forward selection to speed up sparse Gaussian process regression, C. M. Bishop, B. J. Frey, uredniki, *Proceedings of the 9th International Workshop on AI and Statistics*, Key West, Florida, 2003.
- [97] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.-Y. Glorennec, H. Hjalmarsson, A. Juditsky, Nonlinear black-box modelling in system identification: A unified overview, *Automatica*, Vol. 31, No. 12, str. 1691–1724, 1995.
- [98] M. Small, K. Judd, Variable prediction steps and long term prediction, Delovno poročilo, University of Western Australia, Dept. of Mathematics and Statistics, Perth, 1999.
- [99] E. Solak, R. Murray-Smith, W. E. Leithead, D. J. Leith, C. E. Rasmussen, Derivative observations in Gaussian process models of dynamic systems, S. Becker, S. Thrun, K. Obermayer, uredniki, *Proceedings of Advances in Neural Information Processing Systems conference*, Volume 15, str. 529–536, MIT Press, 2003.
- [100] A. Stare, N. Hvala, D. Vrečko, Modelling, identification and validation of models for predictive ammonia control in a wastewater treatment plant — a case study, *ISA Transactions*, Vol. 45, No. 2, str. 159–174, 2006.
- [101] J. Stark, D. S. Broomhead, M. E. Davies, J. Huke, Delay embeddings of forced systems: II Stochastic forcing, *Journal of Nonlinear Science*, Vol. 13, No. 6, str. 519–577, 2003.
- [102] S. Strmčnik, urednik, *Celostni pristop k računalniškemu vodenju procesov*, Založba FE, Ljubljana, 1998.
- [103] J. A. K. Suykens, Support vector machines: A nonlinear modelling and control perspective, *European Journal of Control*, Vol. 7, str. 311–327, 2001.
- [104] J. A. K. Suykens, J. Vandevallé, uredniki, *Nonlinear modelling – advanced black-box techniques*, Kluwer Academic Publishers, Dordrecht, 1998.
- [105] I. Škrjanc, S. Blažič, O. Agamenonni, Identification of dynamical systems with a robust interval fuzzy model, *Automatica*, Vol. 41, str. 327–332, 2005.

- [106] T. Takagi, M. Sugeno, Fuzzy identification of systems and its application to modeling and control, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 15, No. 1, str. 116–132, 1985.
- [107] F. Takens, D. A. Rand, L.-S. Young, *Dynamical Systems and Turbulence*, Volume 898 iz *Lecture Notes in Mathematics*, poglavje Detecting strange attractors in turbulence, str. 366–381, Springer-Verlag, Berlin, 1981.
- [108] D. Temple Lang, Approaches for random number generation, 2004, <http://eeyore.ucdavis.edu/stat141/Notes/RNG.pdf>, pridobljeno 18.12.2006 s svetovnega spleta.
- [109] L. Todorovski, S. Džeroski, Integrating knowledge-driven and data-driven approaches to modelling, *Ecological Modelling*, Vol. 194, str. 3–13, 2006.
- [110] G. Tomšič, T. Slivnik, *Matematika IV*, Fakulteta za elektrotehniko, Ljubljana, 1998.
- [111] V. Venkatasubramanian, R. Rengaswamy, K. Yin, S. N. Kavuri, A review of process fault detection and diagnosis. Part I: Quantitative model-based methods, *Computers and Chemical Engineering*, Vol. 27, str. 293–311, 2003.
- [112] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, A review of process fault detection and diagnosis. Part II: Qualitative models and search strategies, *Computers and Chemical Engineering*, Vol. 27, str. 313–326, 2003.
- [113] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, K. Yin, A review of process fault detection and diagnosis. Part III: Process history based methods, *Computers and Chemical Engineering*, Vol. 27, str. 327–346, 2003.
- [114] D. Vrančič, Dj. Juričič, J. Petrovčič, Measurements and mathematical modelling of a semi-industrial liquid gas separator for the purpose of a fault diagnosis, Delovno poročilo 7260, Institut Jožef Stefan, Ljubljana, 1995.
- [115] L. Wasserman, *All of statistics – A concise course in statistical inference*, Springer texts in statistics, Springer, New York, 2004.
- [116] E. W. Weisstein, Normal Sum Distribution, From MathWorld – A Wolfram Web Resource, <http://mathworld.wolfram.com/NormalSumDistribution.html>, pridobljeno 16.5.2007 s svetovnega spleta.

- 
- [117] E. W. Weisstein, Normal distribution, From MathWorld – A Wolfram Web Resource. <http://mathworld.wolfram.com/NormalDistribution.html>, pridobljeno 16.5.2007 s svetovnega spleta.
- [118] E. W. Weisstein, Toeplitz matrix, From MathWorld – A Wolfram Web Resource. <http://mathworld.wolfram.com/ToeplitzMatrix.html>, pridobljeno 11.6.2007 s svetovnega spleta.
- [119] E. W. Weisstein, Variance, From MathWorld – A Wolfram Web Resource. <http://mathworld.wolfram.com/Variance.html>, pridobljeno 17.5.2007 s svetovnega spleta.
- [120] C. K. I. Williams, Prediction with Gaussian processes: from linear regression and beyond, M.I. Jordan, urednik, *Learning in graphical models*, Volume 89 iz *Nato Science Series D*, str. 599–621, Springer, Berlin, 1998.
- [121] C. K. I. Williams, M. Seeger, Using the Nyström method to speed up kernel machines, T. K. Leen, T. G. Diettrich, V. Tresp, uredniki, *Proceedings of Advances in Neural Information Processing Systems conference*, Volume 13, str. 682–688, MIT Press, MA, 2001.
- [122] C. Zopounidis, P. M. Pardalos, G. Baourakis, *Fuzzy Sets in Management, Economy and Marketing*, World Scientific Publishing Company, 2002.
- [123] *IEEE Transactions on neural networks*, 1990 in naprej.
- [124] *IEEE Transactions on fuzzy systems*, 1993 in naprej.



## Dodatek A

### Nekateri uporabljeni matematični izrazi

#### A.1 Gaussova porazdelitev in nekatere lastnosti

Če je porazdelitvena funkcija naključne spremenljivke  $X$ :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad (\text{A.1})$$

kjer je z  $\mu$  označena njena srednja vrednost in je  $\sigma^2$  njena varianca, pravimo, da je spremenljivka  $X$  porazdeljena po normalni (Gaussovi) porazdelitvi [117] in to označimo kot:

$$X \sim \mathcal{N}(\mu, \sigma^2). \quad (\text{A.2})$$

Nekaj lastnosti:

1. Če je porazdelitev naključne spremenljivke  $X$  normalna (Gaussova)  $X \sim \mathcal{N}(\mu, \sigma^2)$  in za naključno spremenljivko  $Y$  velja  $Y = aX + b$ , kjer sta  $a$  in  $b$  realni števili, potem je tudi naključna spremenljivka  $Y$  porazdeljena normalno s srednjo vrednostjo  $a\mu + b$  in varianco  $(a\sigma)^2$  [116]:

$$Y \sim \mathcal{N}(a\mu + b, (a\sigma)^2). \quad (\text{A.3})$$

2. Naključna spremenljivka  $Y$ , ki jo dobimo kot vsoto dveh naključnih, normalno porazdeljenih spremenljivk  $X_1$  in  $X_2$ ,  $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$  in  $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ , je prav tako porazdeljena normalno [116]:

$$Y = X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2). \quad (\text{A.4})$$

Enako za vsoto  $n$  normalno porazdeljenih naključnih spremenljivk  $X_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$ ,  $n = 1, \dots, N$ , velja, da je porazdeljena normalno [116]:

$$Y = \sum_{n=1}^N X_n \sim \mathcal{N}\left(\sum_{n=1}^N \mu_n, \sum_{n=1}^N \sigma_n^2\right). \quad (\text{A.5})$$

## A.2 Produkt dveh Gaussovih distribucij

Predstavljeni rezultati so povzeti iz [27]. Označimo  $D$ -dimenzionalno Gaussovo porazdelitev s srednjo vrednostjo  $\mathbf{a}$  in kovariančno matriko  $\mathbf{A}$  po  $\mathbf{x}$ -u z  $\mathcal{N}_{\mathbf{x}}(\mathbf{a}, \mathbf{A})$ :

$$\mathcal{N}_{\mathbf{x}}(\mathbf{a}, \mathbf{A}) = (2\pi)^{-D/2} |\mathbf{A}|^{-1/2} \exp \left[ -\frac{1}{2} (\mathbf{x} - \mathbf{a})^T \mathbf{A}^{-1} (\mathbf{x} - \mathbf{a}) \right] \quad (\text{A.6})$$

in enako  $\mathcal{N}_{\mathbf{x}}(\mathbf{b}, \mathbf{B})$ :

$$\mathcal{N}_{\mathbf{x}}(\mathbf{b}, \mathbf{B}) = (2\pi)^{-D/2} |\mathbf{B}|^{-1/2} \exp \left[ -\frac{1}{2} (\mathbf{x} - \mathbf{b})^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{b}) \right]. \quad (\text{A.7})$$

Potem je produkt teh dveh porazdelitev:

$$\mathcal{N}_{\mathbf{x}}(\mathbf{a}, \mathbf{A}) \mathcal{N}_{\mathbf{x}}(\mathbf{b}, \mathbf{B}) = z \mathcal{N}_{\mathbf{x}}(\mathbf{c}, \mathbf{C}) \quad (\text{A.8})$$

kjer

$$\mathbf{c} = \mathbf{C}(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b}), \quad \mathbf{C} = (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}$$

in

$$\begin{aligned} z &= (2\pi)^{-D/2} |\mathbf{C}|^{1/2} |\mathbf{A}|^{-1/2} |\mathbf{B}|^{-1/2} \exp \left[ -\frac{1}{2} (\mathbf{a}^T \mathbf{A}^{-1} \mathbf{a} + \mathbf{b}^T \mathbf{B}^{-1} \mathbf{b} - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{c}) \right] \\ &= (2\pi)^{-D/2} |\mathbf{A} + \mathbf{B}|^{-1/2} \exp \left[ -\frac{1}{2} (\mathbf{a} - \mathbf{b})^T (\mathbf{A} + \mathbf{B})^{-1} (\mathbf{a} - \mathbf{b}) \right] \end{aligned} \quad (\text{A.9})$$

ali tudi:

$$\mathbf{z} = \mathcal{N}_{\mathbf{a}}(\mathbf{b}, \mathbf{A} + \mathbf{B}) \quad (\text{A.10})$$

ali

$$\mathbf{z} = \mathcal{N}_{\mathbf{b}}(\mathbf{a}, \mathbf{A} + \mathbf{B}). \quad (\text{A.11})$$

## A.3 Izrek o invertiranju matrik

(angl. *Matrix Inversion Lemma*) [27]:

Če sta  $\mathbf{A}$  in  $\mathbf{B}$  kvadratni in nesingularni matriki, potem velja:

$$(\mathbf{A} + \mathbf{X}\mathbf{B}\mathbf{X}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{X}(\mathbf{B}^{-1} + \mathbf{X}^T\mathbf{A}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{A}^{-1}. \quad (\text{A.12})$$

## A.4 Statistični momenti

Predstavljeni bosta dva uporabljena statistična momenta, povzeta iz [83,110]. Če je  $x$  zvezna naključna spremenljivka z verjetnostno porazdelitvijo  $p(x)$ , so njeni momenti definirani kot:

- moment prvega reda – srednja vrednost:

$$m = \mathbb{E}[x] = \int_{-\infty}^{+\infty} xp(x)dx, \quad (\text{A.13})$$

- centralni moment drugega reda – varianca:

$$\sigma^2 = \mathbb{E}[(x - m)^2] = \int_{-\infty}^{+\infty} (x - m)^2 p(x)dx. \quad (\text{A.14})$$

Če je  $x$  naključna spremenljivka z vektorjem realizacij  $\mathbf{x}$  dolžine  $N$ , njene momente iz vektorja  $\mathbf{x}$  izračunamo kot:

- moment prvega reda – srednja vrednost:

$$m = \mathbb{E}[x] = \frac{1}{N} \sum_{i=1}^N x_i, \quad (\text{A.15})$$

- centralni moment drugega reda – varianca:

$$\sigma^2 = \mathbb{E}[(x - m)^2] = \frac{1}{N} \sum_{i=1}^N (x_i - m)^2. \quad (\text{A.16})$$

Varianca  $\sigma^2$ , ki jo izračunamo iz (A.16), ni nepristranska. Če porazdelitve, iz katere je realiziran vektor  $\mathbf{x}$ , ne poznamo, je potrebno izraz za varianco ustrezno popraviti, da ta izraža varianco naključne spremenljivke  $x$  [119]:

$$\sigma_{N-1}^2 = \frac{N}{N-1} \sigma^2 \quad (\text{A.17})$$

oz. uporabimo korekcijo:

$$\frac{\sigma_{N-1}^2}{\sigma^2} = \frac{N}{N-1}. \quad (\text{A.18})$$

## A.5 Statistični momenti mešanice Gaussovih porazdelitev

Zanimata nas srednja vrednost in varianca mešanice  $N$  Gaussovih porazdelitev  $\mathcal{N}_n(m_n, v_n)$ ,  $n = 1, \dots, N$ , kjer je vsaka od teh porazdelitev predstavljena s srednjo vrednostjo  $m_n$  in varianco  $v_n$ .

Zaradi lažje izpeljave si predstavljajmo, da je vsaka od Gaussovih porazdelitev  $\mathcal{N}_n$  predstavljena s  $K$  realizacijami naključne spremenljivke, zbranih v vektorju  $\mathbf{x}_n = x_n^1 \dots x_n^k \dots x_n^K$ , kjer indeks  $n$  označuje  $n$ -to porazdelitev,  $k$  pa njeno  $k$ -to realizacijo. Mešanico  $N$  Gaussovih porazdelitev predstavimo z vektorjem realizacij  $\mathbf{x}_0 = [x_1 \dots x_{KN}]^T$ , v katerem združimo realizacije posameznih porazdelitev  $\mathcal{N}_n$ :  $\mathbf{x}_0 = [\mathbf{x}_1 \dots \mathbf{x}_N]^T$ .

Srednjo vrednost  $m_0$  mešanice Gaussovih porazdelitev izračunamo (A.15):

$$m_0 = E[\mathbf{x}_0] = \frac{1}{NK} \sum_{i=1}^{NK} x_i = \frac{1}{N} \sum_{n=1}^N \left( \frac{1}{K} \sum_{k=1}^K x_n^k \right) = \frac{1}{N} \sum_{n=1}^N m_n. \quad (\text{A.19})$$

Varianco  $v_0$  mešanice Gaussovih porazdelitev izračunamo (A.16):

$$\begin{aligned} v_0 &= E[(\mathbf{x}_0 - m_0)^2] \\ &= \frac{1}{NK} \sum_{i=1}^{NK} (x_i - m_0)^2 \\ &= \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K (x_n^k - m_0)^2 \\ &= \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K (x_n^k - m_n + m_n - m_0)^2 \\ &= \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K ((x_n^k - m_n) + (m_n - m_0))^2 \\ &= \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K ((x_n^k - m_n)^2 + 2(x_n^k - m_n)(m_n - m_0) + (m_n - m_0)^2) \\ &= \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K (x_n^k - m_n)^2 + 2 \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K (x_n^k - m_n)(m_n - m_0) + \\ &\quad + \frac{1}{NK} \sum_{n=1}^N \sum_{k=1}^K (m_n - m_0)^2 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N} \sum_{n=1}^N \left( \frac{1}{K} \sum_{k=1}^K (x_n^k - m_n)^2 \right) + 2 \frac{1}{N} \sum_{n=1}^N (m_n - m_0) \overbrace{\left( \frac{1}{K} \sum_{k=1}^K (x_n^k - m_n) \right)}^0 + \\
&\quad + \frac{1}{N} \sum_{n=1}^N (m_n - m_0)^2 \overbrace{\left( \frac{1}{K} \sum_{k=1}^K 1 \right)}^1 \\
&= \frac{1}{N} \sum_{n=1}^N v_n + \frac{1}{N} \sum_{n=1}^N (m_n - m_0)^2
\end{aligned} \tag{A.20}$$

in dobimo:

$$v_0 = \frac{1}{N} \sum_{n=1}^N (v_n + (m_n - m_0)^2). \tag{A.21}$$

Uporabiti moramo še korekcijo (A.18) in dobimo:

$$v_0 = \frac{1}{N-1} \sum_{n=1}^N (v_n + (m_n - m_0)^2). \tag{A.22}$$

## A.6 Toeplitzova matrika

Toeplitzova matrika je matrika, sestavljena iz  $k = 2n - 1$  parametrov  $a_k$ , oblike [118]:

$$\mathbf{A}_T = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots & a_{-n+1} \\ a_1 & a_0 & a_{-1} & \ddots & \vdots \\ a_2 & a_1 & a_0 & \ddots & a_{-2} \\ \vdots & \ddots & \ddots & \ddots & a_{-1} \\ a_{n-1} & \cdots & a_1 & a_1 & a_0 \end{bmatrix}. \tag{A.23}$$

Tipično se uporablja pri reševanju nekaterih oblik diferencialnih enačb, za analizo časovnih vrst, obdelavo signalov itd. [118]

Pri uporabi za konstrukcijo šumnega dela kovariančne matrike (Dodatek C.2) velja:  $a_k = a_{-k}$ ,  $k = 1, \dots, N_a$ , kjer je  $N_a$  red šumnega procesa.



## Dodatek B

### Nekatere izpeljave

#### B.1 Simulacija GP modela z analitično propagacijo negotovosti

Predstavljena bo simulacija GP modela z analitično propagacijo negotovosti, tj. propagacija negotovosti za primer, da integral (3.14) rešimo za (predpostavljene) Gaussove porazdelitve posameznih vhodnih regresorjev. Izpeljava je povzeta iz [27].

#### Napovedovanje z uporabo GP modela pri porazdeljenem vhodu

Denimo, da je vhod v GP model sestavljen iz regresorjev, danih v obliki Gaussove porazdelitve:

$$\mathbf{x} = \boldsymbol{\nu} + \boldsymbol{\epsilon}_x, \quad (\text{B.1})$$

kjer je  $\boldsymbol{\epsilon}_x \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_x)$ . Za napoved pri  $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$  moramo integrirati napovedane porazdelitve  $p(y|\mathcal{D}, \mathbf{x})$  pri natančnem vhodu<sup>1</sup>  $\mathbf{x}$  prek celotne vhodne porazdelitve  $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$ , enačba (3.14). Ker je  $p(y|\mathcal{D}, \mathbf{x})$  nelinearna funkcija  $\mathbf{x}$ -a, nova napovedana porazdelitev  $p(y|\mathcal{D}, \boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$  ni Gaussova in integrala ne moremo rešiti, ne da bi se zatekli k aproksimaciji [27].

Označimo srednjo vrednost in varianco napovedanega izhoda kot  $\mu(\boldsymbol{\nu})$  in  $\sigma^2(\boldsymbol{\nu})$  za primer, ko kot vhod v GP model nastopajo natančne vrednosti  $\boldsymbol{\nu}$ . Ko so vhodi dani kot porazdelitve, sta srednja vrednost in varianca izhodne ne-Gaussove porazdelitve GP modela  $p(y|\mathcal{D}, \boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$  označeni kot:  $m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$ ,  $v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$ . To porazdelitev *aproximiramo* kot Gaussovo z istimi momenti prvega in drugega reda — srednjo vrednostjo in varianco (Dodatek A.4):

$$p(y|\mathcal{D}, \boldsymbol{\nu}, \boldsymbol{\Sigma}_x) \approx \mathcal{N}(m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x), v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)) = N(m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x), v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)). \quad (\text{B.2})$$

<sup>1</sup>Danem v obliki točke in ne porazdelitve.

Iz (3.14) izračunamo srednjo vrednost in varianco porazdelitve (B.2) kot<sup>2</sup>:

$$m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = \int y \left\{ \int p(y|\mathcal{D}, \mathbf{x}) p(\mathbf{x}|\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) d\mathbf{x} \right\} dy \quad (\text{B.3})$$

$$v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = \int y^2 \left\{ \int p(y|\mathcal{D}, \mathbf{x}) p(\mathbf{x}|\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) d\mathbf{x} \right\} dy - m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)^2. \quad (\text{B.4})$$

Z uporabo:

$$\int yp(y|\mathcal{D}, \mathbf{x}) dy = \mu(\mathbf{x}) \quad (\text{B.5})$$

$$\int y^2 p(y|\mathcal{D}, \mathbf{x}) dy = \sigma^2(\mathbf{x}) + \mu^2(\mathbf{x}) \quad (\text{B.6})$$

dobimo

$$m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = \mathbb{E}[\mu(\mathbf{x})] \quad (\text{B.7})$$

$$v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = \mathbb{E}[\sigma^2(\mathbf{x})] + \mathbb{E}[\mu^2(\mathbf{x})] - m^2(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x), \quad (\text{B.8})$$

kjer z  $\mathbb{E}[z(\mathbf{x})] = \int zp(\mathbf{x}|\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) d\mathbf{x}$  označimo povprečje spremenljivke  $z$  v odvisnosti od vhodne porazdelitve  $p(\mathbf{x}|\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$ .

Če v izraza (B.7) in (B.8) vstavimo izraza za  $\mu(\mathbf{x})$  in  $\sigma^2(\mathbf{x})$ , dobimo:

$$\mathbb{E}[\mu(\mathbf{x})] = \sum_{i=1}^N \beta_i \mathbb{E}[C(\mathbf{x}, \mathbf{x}_i)] \quad (\text{B.9})$$

$$\mathbb{E}[\sigma^2(\mathbf{x})] = \mathbb{E}[C(\mathbf{x}, \mathbf{x})] - \sum_{i,j=1}^N K_{ij}^{-1} \mathbb{E}[C(\mathbf{x}, \mathbf{x}_i) C(\mathbf{x}, \mathbf{x}_j)] \quad (\text{B.10})$$

$$\mathbb{E}[\mu^2(\mathbf{x})] = \sum_{i,j=1}^N \beta_i \beta_j \mathbb{E}[C(\mathbf{x}, \mathbf{x}_i) C(\mathbf{x}, \mathbf{x}_j)] \quad (\text{B.11})$$

in za napovedano srednjo vrednost (B.7) in varianco (B.8) dobimo:

$$m(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = \sum_{i=1}^N \beta_i \mathbb{E}[C(\mathbf{x}, \mathbf{x}_i)] \quad (\text{B.12})$$

$$v(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = \mathbb{E}[C(\mathbf{x}, \mathbf{x})] - \sum_{i,j=1}^N (K_{ij}^{-1} - \beta_i \beta_j) \mathbb{E}[C(\mathbf{x}, \mathbf{x}_i) C(\mathbf{x}, \mathbf{x}_j)] - m^2(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x). \quad (\text{B.13})$$

<sup>2</sup>Kadar meje pri računanju integralov v tem dodatku niso dane, predpostavimo integriranje od  $-\infty$  do  $+\infty$ .

Označimo:

$$I = \mathbb{E}[C(\mathbf{x}, \mathbf{x})] = \int C(\mathbf{x}, \mathbf{x})p(\mathbf{x})d\mathbf{x} \quad (\text{B.14})$$

$$I_i = \mathbb{E}[C(\mathbf{x}, \mathbf{x}_i)] = \int C(\mathbf{x}, \mathbf{x}_i)p(\mathbf{x})d\mathbf{x} \quad (\text{B.15})$$

$$I_{ij} = \mathbb{E}[C(\mathbf{x}, \mathbf{x}_i)(\mathbf{x}, \mathbf{x}_j)] = \int C(\mathbf{x}, \mathbf{x}_i)C(\mathbf{x}, \mathbf{x}_j)p(\mathbf{x})d\mathbf{x}, \quad (\text{B.16})$$

kjer je  $p(\mathbf{x})$  porazdelitev vhodne spremenljivke  $\mathbf{x}$ . Kako rešljivi so integrali (B.14)–(B.16), je odvisno od kovariančne funkcije  $C(.,.)$ .

1. Če je kovariančna funkcija linearna, Gaussova, polinomska ali njihova kombinacija, potem so integrali (B.14)–(B.16) izračunljivi analitično. Velja opozoriti, da gre kljub točnosti izračuna integralov še vedno za aproksimacijo, saj izhodna porazdelitev ni Gaussova, ampak jo samo predstavimo kot Gaussovo z enako srednjo vrednostjo in varianco. Napaka, ki jo naredimo z aproksimacijo, je odvisna od sistema in je ne moremo določiti.
2. V primeru drugačne kovariančne funkcije moramo integrale (B.14)–(B.16) aproksimirati, nato uporabimo dobljeno srednjo vrednost in varianco. Nekaj možnosti je pokazanih v [27].

V nadaljevanju bo pokazan izračun integralov (B.14)–(B.16) in iz tega sledeča napoved izhoda GP modela pri porazdeljenem vhodu za Gaussovo kovariančno funkcijo (2.4).

### Predikcija za porazdeljene vhode pri uporabi Gaussove kovariančne funkcije

Vzemimo, da imamo GP model z Gaussovo kovariančno funkcijo (2.4):

$$C(\mathbf{x}_i, \mathbf{x}_j) = C_G(\mathbf{x}_i, \mathbf{x}_j) = v \exp \left[ -\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{W}^{-1}(\mathbf{x}_i - \mathbf{x}_j)^T \right], \quad (\text{B.17})$$

kjer je  $\mathbf{W} = \text{diag}([w_1 \dots w_d]^{-1})$  (alternativni zapis funkcije (2.4)).

Ponovimo rezultat napovedi za natančen vhod brez šuma na izhodu procesa:

$$\mu(\mathbf{x}) = \sum_{i=1}^N \beta_i C_G(\mathbf{x}, \mathbf{x}_i) \quad (\text{B.18})$$

$$\sigma^2(\mathbf{x}) = C_G(\mathbf{x}, \mathbf{x}) - \sum_{i,j=1}^N K_{ij}^{-1} C_G(\mathbf{x}, \mathbf{x}_i) C_G(\mathbf{x}, \mathbf{x}_j), \quad (\text{B.19})$$

kjer velja  $C_G(\boldsymbol{\nu}, \boldsymbol{\nu}) = v$ . Za napoved pri  $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x)$  moramo izračunati:

$$m^G(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = \sum_{i=1}^N \beta_i I_i^G \quad (\text{B.20})$$

$$v^G(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) = I^G - \sum_{i,j=1}^N (K_{ij}^{-1} - \beta_i \beta_j) I_{ij}^G - m^G(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x), \quad (\text{B.21})$$

kjer

$$I^G = \mathbb{E}[C_G(\mathbf{x}, \mathbf{x})] = v \quad (\text{B.22})$$

$$I_i^G = \mathbb{E}[C_G(\mathbf{x}, \mathbf{x}_i)] \quad (\text{B.23})$$

$$I_{ij}^G = \mathbb{E}[C_G(\mathbf{x}, \mathbf{x}_i) C_G(\mathbf{x}, \mathbf{x}_j)]. \quad (\text{B.24})$$

Označimo  $C(\mathbf{x}_i, \mathbf{x}_j) = c N_{\mathbf{x}_i}(\mathbf{x}_j, \mathbf{W})$ , kjer  $c = (2\pi)^{D/2} |\mathbf{W}|^{1/2} v$ . Z uporabo formule za produkt dveh Gaussov (Dodatek A.2) dobimo:

$$I_i^G = c N_{\boldsymbol{\nu}}(\mathbf{x}_i, \mathbf{W} + \boldsymbol{\Sigma}_x) \quad (\text{B.25})$$

$$I_{ij}^G = c^2 N_{\mathbf{x}_i}(\mathbf{x}_j, 2\mathbf{W}) N_{\boldsymbol{\nu}}\left(\frac{\mathbf{x}_i + \mathbf{x}_j}{2}, \boldsymbol{\Sigma}_x + \frac{\mathbf{W}}{2}\right). \quad (\text{B.26})$$

Za napovedano srednjo vrednost in varianco pri Gaussovi kovariančni funkciji in predpostavki o Gaussovi porazdelitvi izhoda tako dobimo:

$$\begin{aligned} m^G(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) &= \sum_{i=1}^N \beta_i c N_{\boldsymbol{\nu}}(\mathbf{x}_i, \mathbf{W} + \boldsymbol{\Sigma}_x) = \\ &= \sum_{i=1}^N \beta_i C_G(\boldsymbol{\nu}, \mathbf{x}_i) C_{\text{mod}_1}(\boldsymbol{\nu}, \mathbf{x}_i) \end{aligned} \quad (\text{B.27})$$

in

$$\begin{aligned} v^G(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) &= C_G(\boldsymbol{\nu}, \boldsymbol{\nu}) - \sum_{i,j=1}^N (K_{ij}^{-1} - \beta_i \beta_j) C_G(\boldsymbol{\nu}, \mathbf{x}_i) C_G(\boldsymbol{\nu}, \mathbf{x}_j) C_{\text{mod}_2}(\boldsymbol{\nu}, \bar{\mathbf{x}}) - \\ &\quad - m^G(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x) \end{aligned} \quad (\text{B.28})$$

kjer  $\bar{\mathbf{x}} = \frac{\mathbf{x}_i + \mathbf{x}_j}{2}$  in

$$C_{\text{mod}_1}(\boldsymbol{\nu}, \mathbf{x}_i) = |\mathbf{I} + \mathbf{W}^{-1} \boldsymbol{\Sigma}_x|^{-1/2} \left[ \frac{1}{2} (\boldsymbol{\nu} - \mathbf{x}_i)^T \Delta^{-1} (\boldsymbol{\nu} - \mathbf{x}_i) \right] \quad (\text{B.29})$$

$$C_{\text{mod}_2}(\boldsymbol{\nu}, \bar{\mathbf{x}}) = \left| \frac{\mathbf{W}^{-1}}{2} \boldsymbol{\Sigma}_x + \mathbf{I} \right|^{-1/2} \exp \left[ \frac{1}{2} (\boldsymbol{\nu} - \bar{\mathbf{x}})^T \Lambda^{-1} (\boldsymbol{\nu} - \bar{\mathbf{x}}) \right], \quad (\text{B.30})$$

kjer  $\Delta^{-1} = \mathbf{W}^{-1} (\mathbf{W}^{-1} + \boldsymbol{\Sigma}_x^{-1})^{-1} \mathbf{W}^{-1}$ , ki se z uporabo pravila o invertiranju matrik (enačba (A.12)) poenostavi v  $\Delta^{-1} = \mathbf{W}^{-1} - (\mathbf{W} + \boldsymbol{\Sigma}_x)^{-1}$ . Enako za  $\Lambda^{-1}$  dobimo:  $\Lambda^{-1} = \left(\frac{\mathbf{W}}{2}\right)^{-1} \left(\left(\frac{\mathbf{W}}{2}\right)^{-1} + \boldsymbol{\Sigma}_x^{-1}\right)^{-1} \left(\frac{\mathbf{W}}{2}\right)^{-1} = 2\mathbf{W}^{-1} - \left(\frac{1}{2}\mathbf{W} + \boldsymbol{\Sigma}_x\right)^{-1}$ .

## Uporaba dobljenih rezultatov za simulacijo dinamičnih sistemov

Pogledali smo si napovedovanje GP modela, pri katerem kot vhod nastopajo porazdelitve. Pokažimo, kako dobljen rezultat apliciramo na simulacijo GP modela. Ideja je povzeta po [27]. Oznaka  $N(.,.)$  namesto  $\mathcal{N}(.,.)$  označuje, da gre samo za aproksimacijo Gaussove porazdelitve.

Kot pri simulaciji GP modela brez propagacije negotovosti (razdelek 3.5.1) predpostavimo, da poznamo vrednosti izhodov iz sistema  $y$  do koraka  $k$ . Zaradi poenostavitve bomo prikazali izpeljavo, kot da ima GP model samo avtoregresijski (AR) del in je torej odvisen samo od zakasnenih vrednosti izhodov. Zakasneni vzorci vhoda, ki predstavljajo del s tekočim povprečjem (MA del), nastopajo kot natančne vrednosti in ne porazdelitve, zato jih bomo dodali na koncu postopka.

Zanima nas predikcija za  $h$  korakov naprej. Velja:

- $k + 1$ :

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

kjer smo 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 (3.12) in (3.13):

$$\mu(\mathbf{x}) = \sum_{i=1}^N \beta_i C(\mathbf{x}_i) \quad (\text{B.31})$$

$$\sigma^2(\mathbf{x}) = C(\mathbf{x}, \mathbf{x}) - \sum_{i,j=1}^N K_{ij}^{-1} C(\mathbf{x}, \mathbf{x}_i) C(\mathbf{x}, \mathbf{x}_j); \quad (\text{B.32})$$

- $k + 2$ :

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

Zdaj kot vhod v model že nastopa predikcija izhoda v prejšnjem koraku simulacije  $\hat{y}(k + 1) \sim \mathcal{N}(\mu(\mathbf{x}_{k+1}), \sigma^2(\mathbf{x}_{k+1}))$ :

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

Srednja vrednost in varianca novega izhoda  $\hat{y}(k + 2)$  se izračunata z uporabo (B.27) in (B.27).

- $k + 3$ :

$$\mathbf{x}_{k+3} = [\hat{y}(k+2), \hat{y}(k+1), \dots, y(k-L+3)]^T,$$

kjer je

$$\hat{y}(k+2) \sim N(m(\mathbf{x}_{k+2}), v(\mathbf{x}_{k+2}))$$

in

$$\hat{y}(k+1) \sim \mathcal{N}(\mu(\mathbf{x}_{k+1}), \sigma^2(\mathbf{x}_{k+1})).$$

Imamo:

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

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 (B.27) in (B.28).

- $\vdots$

- $k + n$ :

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  $\boldsymbol{\Sigma}'_x$ :

$$\begin{aligned} \mathbf{x}_{k+n} &= [\hat{y}(k+n-1), \hat{y}(k+n-2), \dots, \hat{y}(k+n-L)]^T \\ &\sim N(\mathbf{m}_{k+n}, \boldsymbol{\Sigma}'_x), \end{aligned} \quad (\text{B.33})$$

kjer

$$\mathbf{m}_{k+n} = [m(\mathbf{x}_{k+n-1}) \dots m(\mathbf{x}_{k+n-L})]^T \quad (\text{B.34})$$

in

$$\boldsymbol{\Sigma}'_x = \begin{bmatrix} v(x_{k+n-1}) & \text{cov}[\hat{y}(k+n-1), \hat{y}(k+n-2)] & \dots & \text{cov}[\hat{y}(k+n-1), \hat{y}(k+n-L)] \\ \text{cov}[\hat{y}(k+n-2), \hat{y}(k+n-1)] & v(x_{k+n-2}) & \dots & \text{cov}[\hat{y}(k+n-2), \hat{y}(k+n-L)] \\ \vdots & \vdots & \vdots & \vdots \\ \text{cov}[\hat{y}(k+n-L), \hat{y}(k+n-1)] & \text{cov}[\hat{y}(k+n-L), \hat{y}(k+n-2)] & \dots & v(x_{k+n-L}) \end{bmatrix} \quad (\text{B.35})$$

Pripadajoči izhod iz modela s srednjo vrednostjo  $m(\mathbf{x}_{k+n})$  in varianco  $v(\mathbf{x}_{k+n})$  se ponovno izračuna z uporabo (B.27) in (B.28).

•                    ⋮

Ob koraku  $k$  moramo, da bi stopili 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  $\hat{y}(k+n-l)$ ,  $l = 1, \dots, L-1$ , zbranimi v  $\mathbf{x}_{k+n}$ . Ta je dana z izrazom:

$$\begin{aligned} \text{cov}[\hat{y}(k+n), \mathbf{x}_{k+n}] &= \sum_i \beta_i C(\boldsymbol{\nu}_{k+n}, \mathbf{x}_i) C_{\text{mod}_1}(\boldsymbol{\nu}_{k+n}, \mathbf{x}_i) \\ &\quad \left( [\mathbf{I} - \mathbf{W}(\mathbf{W} + \boldsymbol{\Sigma}_{k+n})^{-1}] \mathbf{x}_i - \boldsymbol{\Sigma}_{k+n}(\mathbf{W} + \boldsymbol{\Sigma}_{k+n})^{-1} \boldsymbol{\nu}_{k+n} \right) \end{aligned} \quad (\text{B.36})$$

kjer je  $\boldsymbol{\Sigma}_{k+n} = \boldsymbol{\Sigma}'_x$ .

\*                    \*                    \*

Pokazan je bil postopek simulacije s propagacijo negotovosti za GP model z Gaussovo kovariančno funkcijo. Kot vhodi  $\mathbf{x}_{k+n}$  v GP model so nastopale samo napovedane vrednosti izhoda  $\hat{y}$  v preteklih korakih. Na izhod vplivajo tudi vrednosti vhoda  $u$ , ki so bile v postopku zaradi večje preglednosti izpuščene. Celotni vhodni vektor  $\mathbf{x}_{k+n}$  v koraku  $k+n$  je tako:

$$\begin{aligned} \mathbf{x}_{k+n} &= [\hat{y}(k+n-1), \dots, \hat{y}(k+n-L), u(k+n-1), \dots, u(k+n-L)] \\ &= N(\boldsymbol{\nu}, \boldsymbol{\Sigma}_x), \end{aligned} \quad (\text{B.37})$$

kjer sta  $\boldsymbol{\nu}$  in  $\boldsymbol{\Sigma}_x$ :

$$\boldsymbol{\nu} = [m(\mathbf{x}_{k+n-1}), \dots, m(\mathbf{x}_{k+n-L}), u(k+n-1), \dots, u(k+n-L)]^T \quad (\text{B.38})$$

in

$$\boldsymbol{\Sigma}_x = \begin{bmatrix} \boldsymbol{\Sigma}'_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (\text{B.39})$$

Prikazani postopek simulacije upošteva vpliv variance napovedanih izhodov na obnašanje sistema. Šele na koncu postopka varianci dobljenih izhodov GP modela prištejemo varianco šuma na izhodu  $v_0$ , saj ta ne vpliva na sistem, ampak zgolj moti vrednost njegovega izhoda.

## B.2 Simulacija FSGP modela

### Napovedovanje z uporabo FSGP modela

Predpostavimo, da poznamo vrednosti vhodnega signala  $u(t)$ , izhoda sistema  $y(t)$  in vrednosti vektorja razvrstilnih spremenljivk  $\boldsymbol{\rho}(t)$  do časa  $t_0$ . Prav tako za začetek predpostavimo, da natančno poznamo funkcijo, ki opisuje odvisnost spremenljivih (tj. nekonstantnih) parametrov LPV modela  $\boldsymbol{\theta}(\boldsymbol{\rho}(t))$  od vrednosti razvrstilnih spremenljivk  $\boldsymbol{\rho}(t)$ .

Enačbi (4.32) in (4.33) lahko pri zapisu sistema v kanonični obliki (4.40) zapišemo v obliki:

$$\dot{y}(t) = \boldsymbol{\theta}(t - T) \mathbf{z}(t - T), \quad (\text{B.40})$$

kjer smo z  $\mathbf{z}(t - T)$  označili odvode regresorjev modela, s  $\boldsymbol{\theta}(t - T)$  pa parametre lokalnih modelov, odvisne od vrednosti razvrstilnih spremenljivk  $\boldsymbol{\rho} = \boldsymbol{\rho}(t - T)$ :

$$\boldsymbol{\theta}(t - T) = [-a_1(t - T) \cdots -a_L(t - T) \ b_1(t - T) \cdots b_L(t - T)] \quad (\text{B.41})$$

$$\mathbf{z}(t - T) = [\dot{y}(t - T) \cdots \dot{y}(t - LT) \ \dot{u}(t - T) \cdots \dot{u}(t - LT)]^T. \quad (\text{B.42})$$

Izhod  $y$  iz sistema v času  $t_0 + \Delta$ ,  $0 < \Delta \leq T$  lahko zapišemo kot:

$$\begin{aligned} y(t_0 + \Delta) &= \int_0^{t_0 + \Delta} \dot{y}(t) dt \\ &= y(t_0) + \int_{t_0}^{t_0 + \Delta} \dot{y}(t) dt. \end{aligned} \quad (\text{B.43})$$

Vrednost odvoda izhoda  $\dot{y}$  v času  $t$  je dana z izrazom (B.40). Če ga vstavimo v enačbo (B.43), za spremembo izhoda  $y$  od časa  $t_0$  do  $t_0 + \Delta$  dobimo:

$$\begin{aligned} y(t_0 + \Delta) - y(t_0) &= \int_{t_0}^{t_0 + \Delta} \dot{y}(t) dt \\ &= \int_{t_0}^{t_0 + \Delta} \boldsymbol{\theta}(t - T) \mathbf{z}(t - T) dt. \end{aligned} \quad (\text{B.44})$$

V primeru, da bi bila odvisnost parametrov lokalnih modelov  $\boldsymbol{\theta} = \boldsymbol{\theta}(\boldsymbol{\rho}(t - T))$  od vektorja razvrstilnih spremenljivk  $\boldsymbol{\rho}(t - T)$  znana, bi bilo iz enačbe (B.44) enostavno oceniti vrednost izhoda  $y(t)$  za  $t_0 < t \leq t_0 + T$ . Vendar točne parametre lokalnih modelov poznamo samo v središčih<sup>3</sup> identificiranih lokalnih modelov, v preostalih točkah, določenih z  $\boldsymbol{\rho}(t - T)$ , pa vrednosti parametrov dobimo z uporabo GP modelov.

<sup>3</sup>Kot pri LMGP modelu tudi tu pod središčem lokalnega modela razumemo točko, okoli katere je bil lokalni model identificiran, določeno z vektorjem razvrstilnih spremenljivk  $\boldsymbol{\rho}$ .

Označimo napoved  $i$ -tega parametra  $\theta_i(t - T)$  iz vektorja parametrov  $\boldsymbol{\theta}(t - T)$  z  $\mathcal{N}_{\theta_i}(\mu_{\theta_i}, \sigma_{\theta_i}^2, t - T)$ ,  $i$ -ti odvod regresorja iz vektorja  $\mathbf{z}(t - T)$  pa z  $z_i(t - T)$ .

Odvod signala  $\dot{y}(t)$ , iz katerega bomo izračunali prispevek k spremembi izhoda od časa  $t$  do  $t + dt$ , lahko izračunamo iz enačbe (B.40):

$$\hat{y}(t) = \boldsymbol{\theta}(t - T) \mathbf{z}(t - T) \quad (\text{B.45})$$

$$= \sum_{i=1}^{2L} z_i(t - T) \mathcal{N}_{\theta_i}(\mu_{\theta_i}, \sigma_{\theta_i}^2, t - T). \quad (\text{B.46})$$

Odводи regresorjev  $z_i$ ,  $i = 1, \dots, 2L$ , so skalarji, zato je vsak člen vsote  $\left( z_i \mathcal{N}_i(\mu_{\theta_i}, \sigma_{\theta_i}^2, t - T) \right)$  iz enačbe (B.46) tudi Gaussova porazdelitev, enačba (A.3). Ker je vsota neodvisnih<sup>4</sup> normalnih porazdelitev tudi normalna porazdelitev, enačba (A.5), lahko oceno odvoda signala v času  $t$  iz enačbe (B.46) zapišemo kot:

$$\hat{y}(t) = \sum_{i=1}^{2L} z_i(t - T) \mathcal{N}_i(\mu_{\theta_i}, \sigma_{\theta_i}^2, t - T) \quad (\text{B.47})$$

$$\stackrel{(A.3)}{=} \sum_{i=1}^{2L} \mathcal{N}_i\left(z_i(t - T)\mu_{\theta_i}(t - T), z_i^2(t - T)\sigma_{\theta_i}^2(t - T)\right) \quad (\text{B.48})$$

$$\stackrel{(A.5)}{=} \mathcal{N}\left(\sum_{i=1}^{2L} z_i(t - T)\mu_{\theta_i}(t - T), \sum_{i=1}^{2L} z_i^2(t - T)\sigma_{\theta_i}^2(t - T)\right) \quad (\text{B.49})$$

$$= \mathcal{N}\left(\mu_{\dot{y}}(t), \sigma_{\dot{y}}^2(t)\right). \quad (\text{B.50})$$

Za izhod sistema v času  $t_0 + \Delta$  iz enačbe (B.44) z uporabo (B.50) dobimo:

$$\hat{y}(t_0 + \Delta) = y(t_0) + \int_{t_0}^{t_0 + \Delta} \mathcal{N}\left(\mu_{\dot{y}}(t), \sigma_{\dot{y}}^2(t)\right) dt \quad (\text{B.51})$$

$$\stackrel{(A.5)}{=} y(t_0) + \mathcal{N}\left(\int_{t_0}^{t_0 + \Delta} \mu_{\dot{y}}(t) dt, \int_{t_0}^{t_0 + \Delta} \sigma_{\dot{y}}^2(t) dt\right) \quad (\text{B.52})$$

$$= \mathcal{N}\left(y(t_0) + \int_{t_0}^{t_0 + \Delta} \mu_{\dot{y}}(t) dt, \int_{t_0}^{t_0 + \Delta} \sigma_{\dot{y}}^2(t) dt\right) \quad (\text{B.53})$$

oz. zapisano drugače:

$$\mu_y(t_0 + \Delta) = y(t_0) + \int_{t_0}^{t_0 + \Delta} \mu_{\dot{y}}(t) dt \quad (\text{B.54})$$

$$\sigma_y^2(t_0 + \Delta) = \int_{t_0}^{t_0 + \Delta} \sigma_{\dot{y}}^2(t) dt. \quad (\text{B.55})$$

<sup>4</sup>Predpostavljamo, da so porazdelitve  $\mathcal{N}_i(\mu_{\theta_i}, \sigma_{\theta_i}^2, t - T)$ ,  $i = 1, \dots, 2L$ , med seboj neodvisne.

Z zveznim premikom  $\Delta$  od 0 do  $T$  tako lahko ocenimo vse vrednosti izhoda  $y$  za čase od  $t_0$  do  $t_0 + T$ . Za oceno izhoda pri času  $t_0 + T$  dobimo:

$$\hat{y}(t_0 + T) = \mathcal{N}(\mu_y(t_0 + T), \sigma_y^2(t_0 + T)) \quad (\text{B.56})$$

$$\mu_y(t_0 + T) = y(t_0) + \int_{t_0}^{t_0 + T} \mu_{\dot{y}}(t) dt \quad (\text{B.57})$$

$$\sigma_y^2(t_0 + T) = \int_{t_0}^{t_0 + T} \sigma_{\dot{y}}^2(t) dt, \quad (\text{B.58})$$

kjer sta

$$\mu_{\dot{y}}(t) = \sum_{i=1}^{2L} z_i(t - T) \mu_{\theta_i}(t - T) \quad (\text{B.59})$$

$$\sigma_{\dot{y}}^2(t) = \sum_{i=1}^{2L} z_i^2(t - T) \sigma_{\theta_i}^2(t - T). \quad (\text{B.60})$$

### Simulacija FSGP modela brez propagacije negotovosti

V prejšnjem razdelku smo opisali napovedovanje izhoda FSGP modela  $\hat{y}(t_0 + T) = \mathcal{N}(\mu_y(t_0 + T), \sigma_y^2(t_0 + T))$ , če poznamo vse potrebne vrednosti odvodov signalov  $\dot{y}(t)$ ,  $\dot{u}(t)$  in vrednosti razvrstilnih spremenljivk  $\boldsymbol{\rho}(t)$  do časa  $t \leq t_0$ :

$$\mu_y(t_0 + T) = y(t_0) + \int_{t_0}^{t_0 + T} \left( \sum_{i=1}^{2L} z_i(t - T) \mu_{\theta_i}(t - T) \right) dt \quad (\text{B.61})$$

$$\sigma_y^2(t_0 + T) = \int_{t_0}^{t_0 + T} \left( \sum_{i=1}^{2L} z_i^2(t - T) \sigma_{\theta_i}^2(t - T) \right) dt. \quad (\text{B.62})$$

Podobno lahko izhod sistema v času  $t_0 + T + \Delta$  za poljuben  $\Delta \geq 0$  napovemo, če poznamo vrednosti odvodov signalov  $\dot{y}$ ,  $\dot{u}$  in razvrstilnih spremenljivk  $\boldsymbol{\rho}$  do časa  $t \leq t_0 + \Delta$ :

$$\mu_y(t_0 + T + \Delta) = y(t_0) + \int_{t_0}^{t_0 + T + \Delta} \left( \sum_{i=1}^{2L} z_i(t - T + \Delta) \mu_{\theta_i}(t - T + \Delta) \right) dt \quad (\text{B.63})$$

$$\sigma_y^2(t_0 + T + \Delta) = \int_{t_0}^{t_0 + T + \Delta} \left( \sum_{i=1}^{2L} z_i^2(t - T + \Delta) \sigma_{\theta_i}^2(t - T + \Delta) \right) dt. \quad (\text{B.64})$$

Pri tem smo za napovedi odvodov signalov  $\mathbf{z}(t)$  in razvrstilnih spremenljivk  $\boldsymbol{\rho}(t)$  uporabili najbolj verjetne vrednosti in ne njihovih porazdelitev.

## Dodatek C

### Algoritmi za delo z GP modeli

#### C.1 Simulacija GP modela z numerično propagacijo negotovosti

Prikazujemo algoritem drugega postopka simulacije GP modela z numerično propagacijo negotovosti (str. 43) v obliki psevdokode.

$[\mathbf{m}, \mathbf{v}] = \text{singpnumer}(\mathcal{D}, \Theta, \mathbf{u}, \mathbf{u}_0, \mathbf{y}_0, S)$

*Vhodi:*

$\mathcal{D}$  ... učna množica GP modela,  $\mathcal{D} = \mathbf{X}, \mathbf{y}$ ;

$\Theta$  ... vrednosti hiperparametrov;

$\mathbf{u}$  ... vzorčen vhodni signal (vektor dolžine  $K$ );

$\mathbf{u}_0, \mathbf{y}_0$  ... začetne vrednosti vhodov in izhodov, tj. vrednosti pred nastopom simulacije, dolžine  $L$  za izhod  $y$  in dolžine  $L - 1$  za vhode  $u$ , kjer je  $L$  red modela;

$S$  ... število ponovitev simulacije (število vzorcev porazdelitve, ki jih uporabimo za njeno aproksimacijo.)

*Izhoda:*

$\mathbf{m}$  ... vektor srednjih vrednosti napovedanega izhoda simuliranega modela dolžine  $K$  in

$\mathbf{v}$  ... vektor varianc napovedanega izhoda simuliranega modela dolžine  $K$

*Oznake:*

$\mathbf{x}_k$  ... vhod v GP model v  $k$ -tem koraku v obliki

$[y(k-1) \dots y(k-L) \ u(k-1) \dots u(k-L)]$ ;

$[m_k, v_k] = \text{GP}(\mathcal{D}, \Theta, \mathbf{x}_k)$  ... napovedovanje izhoda GP modela, danega z  $\mathcal{D}$  in

$\Theta$  v  $k$ -tem koraku pri vходу v model  $\mathbf{x}_k$ ;

$y_{k-1}^s$  ... naključen vzorec iz napovedane porazdelitve  $N(m_{k-1}, v_{k-1})$ .

Postopek:

1. (Op. Inicializacija matrik  $\mathbf{M}$  in  $\mathbf{V}$ , kamor bomo vpisovali rezultate – srednjo vrednost in varianco – posameznih realizacij simulacije  $s$  v vsakem koraku  $k$ )

$$\mathbf{M} = \text{zeros}(S, K)$$

$$\mathbf{V} = \text{zeros}(S, K)$$

2. zanka po  $s = 1 \dots S$

- korak 1:

$$\mathbf{x}_1 = [\mathbf{y}_0 \ u(1) \ \mathbf{u}_0]$$

$$[m_1, v_1] = \text{GP}(\mathcal{D}, \Theta, \mathbf{x}_1)$$

$$\mathbf{M}(s, 1) = m_1$$

$$\mathbf{V}(s, 1) = v_1$$

- korak 2:

vzorčimo  $y_1^s$  iz  $N(m_1, v_1)^1$

$$\mathbf{x}_2 = [y_1^s \ \mathbf{y}_0(1 : L - 1) \ u(2) \ u(1) \ \mathbf{u}_0(1 : L - 2)]$$

$$[m_2, v_2] = \text{GP}(\mathcal{D}, \Theta, \mathbf{x}_2)$$

$$\mathbf{M}(s, 2) = m_2$$

$$\mathbf{V}(s, 2) = v_2$$

- korak 3:

vzorčimo  $y_2^s$  iz  $N(m_2, v_2)$

$$\mathbf{x}_3 = [y_2^s \ y_1^s \ \mathbf{y}_0(1 : L - 2) \ u(3) \ u(2) \ u(1) \ \mathbf{u}_0(1 : L - 3)]$$

$$[m_3, v_3] = \text{GP}(\mathcal{D}, \Theta, \mathbf{x}_3)$$

$$\mathbf{M}(s, 3) = m_3$$

$$\mathbf{V}(s, 3) = v_3$$

•  $\vdots$

- korak  $k$ :

vzorčimo  $y_{k-1}^s$  iz  $N(m_{k-1}, v_{k-1})$

$$\mathbf{x}_k = [y_{k-1}^s \dots y_{k-L}^s \ u(k-1) \dots u(k-L)]$$

$$[m_k, v_k] = \text{GP}(\mathcal{D}, \Theta, \mathbf{x}_k)$$

$$\mathbf{M}(s, k) = m_k$$

$$\mathbf{V}(s, k) = v_k$$

•  $\vdots$

- korak  $K$ :

vzorčimo  $y_{K-1}^s$  iz  $N(m_{K-1}, v_{K-1})$

---

<sup>1</sup>V našem primeru smo uporabljali kar vgrajeno Matlabovo funkcijo za naključno vzorčenje *randn* iz Gaussove porazdelitve (metoda Ziggurat [68]), lahko pa bi naključno vzorčenje izvedli npr. z metodo z invertiranjem kumulativne porazdelitvene funkcije ali metodo MCMC [108].

$$\begin{aligned}\mathbf{x}_K &= [y_{K-1}^s \dots y_{K-L}^s u(K-1) \dots u(K-L)] \\ [m_K, v_K] &= \text{GP}(\mathcal{D}, \boldsymbol{\Theta}, \mathbf{x}_K) \\ \mathbf{M}(s, K) &= m_K \\ \mathbf{V}(s, K) &= v_K\end{aligned}$$

3. (Op. V prvem koraku simulacije natančno poznamo vhod, zato so vsi rezultati enaki:)

$$m(1) = M(1, 1)$$

$$v(1) = V(1, 1)$$

(Op. Iz mešanice  $S$  Gaussovih porazdelitev, dobljenih s simulacijo, za vsak korak  $k$ ,  $k = 2, \dots, K$ , izračunamo srednjo vrednost  $m(k)$  in varianco  $v(k)$  skupne porazdelitve, prim. Dodatek A.5)

za  $k = 2 \dots K$ :

$$m(k) = \frac{1}{S}(\mathbf{M}(:, k)) \text{ in}$$

$$v(k) = \frac{1}{S-1}(\mathbf{V}(:, k)) + \frac{1}{S-1} \left( \sum_{s=1}^S (M(s, k) - m(k))^2 \right)$$

4. vrnemo vektorja srednjih vrednosti in varianc  $\mathbf{m}$  in  $\mathbf{v}$ .  
(Konec postopka)

Opisani algoritem opisuje metodo simulacije z numerično propagacijo negotovosti, v disertaciji označen kot *drugi postopek* (str. 43). Če bi želeli uporabiti *prvi postopek*, se postopek bistveno ne spremeni, le v vsakem koraku simulacije  $k$  (že vzorčene) vrednosti izhoda  $y_{k-l}^s$  za  $l = 2, \dots, L$  vzorčimo ponovno.

## C.2 Konstrukcija šumnega dela kovariančne matrike za modeliranje barvnega šuma na izhodu sistema

Prikazujemo algoritem v obliki psevdokode za konstrukcijo šumnega dela kovariančne matrike  $\mathbf{K}_n$ , če poznamo model šumnega procesa na izhodu sistema:

- ARMA model, določen z vektorjema  $\mathbf{a} = [1 \ a_1 \ \dots \ a_r]$  in  $\mathbf{b} = [b_1 \ \dots \ b_r]$ , ter
- varianco  $\sigma^2$ .

$[\mathbf{K}_n, \mathbf{k}_n] = \text{covNoiseClr}(N, \mathbf{a}, \mathbf{b}, \sigma^2, \mathbf{ix})$

*Vhodi:*

$N$  ... maksimalna razlika med dvema učnima vzorcema kovariančna matrike v korakih oz. (velikost kovariančne matrike-1);

$\mathbf{a}$  ... vodoravni vektor, ki opisuje AR-del prenosne funkcije šuma na izhodu;

$\mathbf{b}$  ... vodoravni vektor, ki opisuje MA-del prenosne funkcije šuma na izhodu;

$\sigma^2$  ... varianca šuma;

$\mathbf{ix}$  ... vektor dolžine  $N$ , v katerem so zapisane razlike med prvim vzorcem in ostalimi vzorci, privzeto:  $\mathbf{ix} = [0 \ 1 \ \dots \ N - 1]$ .

*Izhoda:*

$\mathbf{K}_n$  ... kovariančna matrika in

$\mathbf{k}_n$  ... vektor kovarianc.

*Postopek:*

1. preveri, če obstaja vhod  $\mathbf{ix}$ :  
da: uporabi in  $N = \max(\mathbf{ix})$ ,  
ne: uporabi privzetega;
2. iz vektorjev  $\mathbf{a}$  in  $\mathbf{b}$  določi reda *AR* in *MA* dela šumnega procesa  
 $N_a = \text{length}(\mathbf{a}) - 1$ ;  $N_b = \text{length}(\mathbf{b}) - 1$   $r = \max(N_a, N_b)$   
če  $N_a \neq N_b$ : dopolni krajši vektor z ničlami;
3. preveri, če je  $N$  večje od reda šuma  
če  $N < N_a$ : končaj postopek.
4. *Konstrukcija MA dela:*  
(Op. Kovarianco zaradi MA dela bomo opisali z vektorjem  $\mathbf{ma}$ )  
če  $N_b = 0$ :  
 $\mathbf{ma} = [\sigma^2 \ b_0^2, \text{zeros}(1, r)]$ ,  
sicer: za  $i = 0, \dots, r$ :  
 $\mathbf{ma}(i) = \sigma^2 (\mathbf{b}(0 : r - i) \cdot \mathbf{b}(i : r))$

## 5. Konstrukcija AR dela:

če  $N_a > 0$ :(Op. Konstrukcija pomožne matrike  $\mathbf{A}_{in}$  Toeplitzove oblike (A.23)) $\mathbf{A}_{in} = \text{toeplitz}(0 : N_a)$ (Op. Tvorjenje Yule-Walkerjeve matrike  $\mathbf{A}$ ) $\mathbf{A} = \text{zeros}(N_a + 1, N_a + 1)$ za  $i = 0 : N_a$ za  $j = 0 : N_a$  $s = A_{in}(i, j)$  $A(i, s) = A(i, s) + a(j);$ 

end

end

(Op. Yule-Walkerjev sistem enačb)

 $\mathbf{A} \cdot \mathbf{k}_n = \mathbf{ma}$ ,kjer je  $\mathbf{k}_n = [C(0)C(1) \dots C(r)]$  vektor kovarianc zaradi šuma $\mathbf{k}_n = \text{inv}(\mathbf{A}) \cdot \mathbf{ma}^T$ 

(Op. Še ostali elementi kovariančnega vektorja:)

za  $i = N_a + 1 : N$  $k_n(i) = -\sum (\mathbf{a}(2 : \text{end}) \cdot \mathbf{k}_n(i - 1 : -1 : i - N_a))$ 

end

če  $N_a = 0$ : $\mathbf{k}_n = [\mathbf{ma}, \text{zeros}(1, N - \text{length}(\mathbf{ma}))]$ 6. (Op. V vektorju  $\mathbf{k}_n$  so spravljene vse potrebne vrednosti kovarianc zaradi šumnega procesa, v šumni del kovariančne matrike jih zlagamo glede na razlike med indeksi posameznih vhodov, spravljenimi v vektorju  $\mathbf{ix}$ )če  $\mathbf{ix} == [0 \ 1 \ \dots \ N - 1]$ : $\mathbf{K}_n = \text{toeplitz}(\mathbf{k}_n)$ sicer:  $l_{ix} = \text{length}(\mathbf{ix})$ za  $i = 1 : l_{ix}$  $\mathbf{p} = \mathbf{ix}(i : l_{ix}) - \mathbf{ix}(i)$  $\mathbf{K}_n(1, i : l_{ix}) = \mathbf{k}_n(\mathbf{p})$  $\mathbf{K}_n(i : l_{ix}, 1) = \mathbf{k}_n(\mathbf{p})^T$ 7. vrni kovariančno matriko  $\mathbf{K}_n$ 

(Konec postopka)

### C.3 Simulacija FSGP modela

Prikazujemo algoritem funkcije, ki je podlaga za simulacijo FSGP modela, za klicanje iz okolja Matlab/Simulink v psevdokodi. Identificirani FSGP model je  $L$ -tega reda.

$$[\hat{y}, \boldsymbol{\mu}_\theta, \boldsymbol{\sigma}_\theta^2] = \text{fsgpmodelfun}(\boldsymbol{\rho})$$

*Vhodi:*

$\boldsymbol{\rho}$  ... vektor razvrstilnih spremenljivk  $\boldsymbol{\rho}(t-T)$ , navadno, ne pa nujno, podmnožica zakasnenih signalov:  $[y(t-T) \dots y(t-LT) \ u(t-T) \dots u(t-LT)]$ ;

$\mathbf{z}$  ... vektor odvodov zakasnenih signalov

$$[\dot{y}(t-T) \dots \dot{y}(t-LT) \ \dot{u}(t-T) \dots \dot{u}(t-LT)];$$

(Op. "Spremenljivki"  $\mathcal{D}_i$  in  $\boldsymbol{\Theta}_i$  sta v FSGP modelu konstantni, zato ju ni potrebno podajati kot parameter funkciji, ampak ju lahko definiramo kot globalni spremenljivki)

$\mathcal{D}_i$  ... učne množice GP modelov,  $i = 1, \dots, 2L$ ;

$\boldsymbol{\Theta}_i$  ... vrednosti hiperparametrov GP modelov,  $i = 1, \dots, 2L$ .

*Izhodi:*

$\hat{y}$  ... napoved odvoda izhoda modela  $\hat{y}(t)$ ; za napoved izhoda  $\hat{y}$  ga je potrebno integrirati;

$\boldsymbol{\mu}_\theta$  ... vektor napovedanih srednjih vrednosti parametrov FSGP modela  $\boldsymbol{\mu}_\theta(\boldsymbol{\rho})$ ;

$\boldsymbol{\sigma}_\theta^2$  ... vektor napovedanih varianc parametrov FSGP modela  $\boldsymbol{\sigma}_\theta^2(\boldsymbol{\rho})$ .

*Oznake:*

$z_i$  ...  $i$ -ti odvod iz vektorja odvodov signalov  $\mathbf{z}$ ,

$[\mu_{\theta i}, \sigma_{\theta i}^2] = \text{GP}(\mathcal{D}_i, \boldsymbol{\Theta}_i, \boldsymbol{\rho})$  ... napovedovanje  $i$ -tega parametra FSGP modela glede na vrednosti razvrstilnih spremenljivk  $\boldsymbol{\rho}$ .

*Postopek:*

1. (Op. Izračunamo vrednosti parametrov modela, odvisne od vektorja razvrstilnih spremenljivk)  
za  $i = 1, \dots, 2L$   
 $[\mu_{\theta i}, \sigma_{\theta i}^2] = \text{GP}(\mathcal{D}_i, \boldsymbol{\Theta}_i, \boldsymbol{\rho})$
2. (Op. Izračunamo odvod napovedanega izhoda iz modela)  
 $\hat{y} = \sum_{i=1}^{2L} \mu_{\theta i} z_i$
3. vrnemo  $\hat{y}$ ,  $\boldsymbol{\mu}_\theta$  in  $\boldsymbol{\sigma}_\theta^2$   
(Konec postopka)

## C.4 Napovedovanje parametrov FSGP modela

Opis funkcije, s katero računamo napovedane parametre FSGP modela  $\hat{\theta}$ , za klicanje iz okolja Matlab/Simulink. Identificirani FSGP model je  $L$ -tega reda.

$$[\boldsymbol{\mu}_\theta, \boldsymbol{\sigma}_\theta^2] = \text{getfsgpparams}(\boldsymbol{\rho})$$

*Vhodi:*

$\boldsymbol{\rho}$  ... vektor razvrstilnih spremenljivk  $\boldsymbol{\rho}(t-T)$ , navadno, ne pa nujno, podmnožica zakasnenih signalov:  $[y(t-T) \dots y(t-LT) \ u(t-T) \dots u(t-LT)]$ ;

(Op. "Spremenljivki"  $\mathcal{D}_i$  in  $\boldsymbol{\Theta}_i$  sta v FSGP modelu konstantni, zato ju ni potrebno podajati kot parameter funkciji, ampak ju lahko definiramo kot globalni spremenljivki)

$\mathcal{D}_i$  ... učne množice GP modelov,  $i = 1, \dots, 2L$ ;

$\boldsymbol{\Theta}_i$  ... vrednosti hiperparametrov GP modelov,  $i = 1, \dots, 2L$ .

*Izhodi:*

$\boldsymbol{\mu}_\theta$  ... vektor napovedanih srednjih vrednosti parametrov FSGP modela  $\boldsymbol{\mu}_\theta(\boldsymbol{\rho})$ ,

$\boldsymbol{\sigma}_\theta^2$  ... vektor napovedanih varianc parametrov FSGP modela  $\boldsymbol{\sigma}_\theta^2(\boldsymbol{\rho})$ .

*Oznake:*

$[\mu_{\theta i}, \sigma_{\theta i}^2] = \text{GP}(\mathcal{D}_i, \boldsymbol{\Theta}_i, \boldsymbol{\rho})$  ... napovedovanje  $i$ -tega parametra FSGP modela z ustreznim GP modelom glede na vrednosti razvrstilnih spremenljivk  $\boldsymbol{\rho}$ .

*Postopek:*

1. (Op. Izračunamo vrednosti parametrov modela, odvisne od vektorja razvrstilnih spremenljivk)
  - za  $i = 1, \dots, 2L$ 

$$[\mu_{\theta i}, \sigma_{\theta i}^2] = \text{GP}(\mathcal{D}_i, \boldsymbol{\Theta}_i, \boldsymbol{\rho})$$
2. vrnemo  $\boldsymbol{\mu}_\theta$  in  $\boldsymbol{\sigma}_\theta^2$   
(Konec postopka)



## Dodatek D

### Interpretacija Takensovega teorema

Leith in Leithead sta pokazala [56], da je za točen opis dinamike nelinearnega dinamičnega sistema potrebna izvedba modela z redundandnim modelom. Takensov teorem (angl. *Takens' embedding theorem*) za vzbujujane sisteme [101] pove, kako z analizo časovnih vrst<sup>1</sup> rekonstruiramo model zveznega dinamičnega sistema. Zadostni pogoj za rekonstrukcijo sistema je:

$$L \geq 2d + 1, \tag{D.1}$$

kjer je  $L$  red diskretnega modela in  $d$  red originalnega zveznega sistema, vendar se v praksi izkaže, da je marsikdaj dovolj že model manjšega reda [32], kar je odvisno od samega sistema:

$$d \leq L \leq 2d + 1. \tag{D.2}$$

Takensov teorem ima dve pomembni posledici:

1. Diskretni model zveznega nelinearnega sistema, identificiran iz vhodno-izhodnih signalov, mnogokrat zadovoljivo opisuje nelinearni sistem le, če je višjega reda, kot je red teoretičnega zveznega modela.
2. Višji red modela, s katerim želimo identificirati nelinearen dinamični sistem poveča računsko breme in lahko poudari “prekletstvo dimenzije” (str. 21).

Velja omeniti, da je Takens predlagal, da za opis nelinearnega dinamičnega sistema uporabljamo po  $L$  zaporednih vzorcev signala, npr.  $\mathbf{x}_k = [x_{k-1} \dots x_{k-L}]$ ; iz teh hočemo izluščiti informacijo o sistemu, kar je v nekaterih vnaprejšnjih strukturah, npr. vnaprejšnja ANN, GP model itd., enostavno izvesti [86].

Takensov teorem se nahaja v [107].

---

<sup>1</sup>Predpostavljamo, da je sistem opisan z vhodnim in izhodnim signalom ter da nimamo vpogleda v skrita stanja sistema.



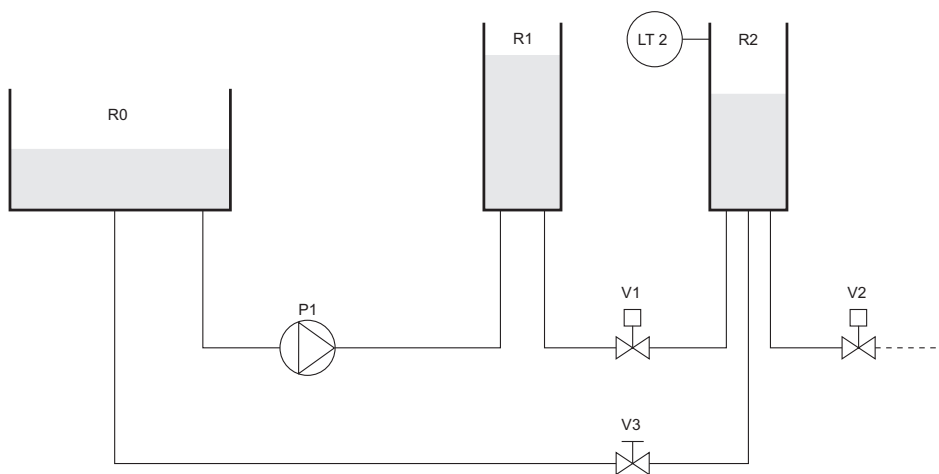
## Dodatek E

### Opisi procesov

#### E.1 Hidravlična modelna naprava

V tem dodatku bo opisana hidravlična modelna naprava, na kateri je bilo predstavljeno vodenje na podlagi FSGP modela (podpoglavje 5.4) in delovanje algoritma za zaznavanje napak na podlagi LMGP modela (podpoglavje 5.6).

Hidravlična modelna naprava, imenovana tudi modelna naprava treh posod [19,89], je nelinearni dinamični sistem, namenjen preizkušanju metod modeliranja in vodenja na Odseku za sisteme in vodenje na Institutu Jožef Stefan. Naprava omogoča izvedbo sistema tretjega reda, vendar so pri taki izvedbi meritve izvedljive v preozkem področju, zato se nelinearnost ne more izraziti. Glede na to je uporabljen podsistem naprave, prikazan na sliki E.1.



**Slika E.1:** Procesna shema uporabljenega podsistema

Podsistem sestavljata posodi R1 in R2, ki sta prek cevi povezana s shranjevalnikom tekočin R0. Nivoja tekočine  $h_1$  in  $h_2$  v posodah R1 in R2 sta

spremenljivki stanj sistema [19, 89] in sistem lahko opišemo z modelom drugega reda. Za črpanje iz shranjevalnika ima posoda R1 v cevi vgrajeno črpalko P1, ki jo poganja enosmerni motor s trajnim magnetom. Hitrost vrtenja motorja  $\omega_1$  je regulirana z analognim regulatorjem. Časovna konstanta spremembe hitrosti vrtenja motorja je zelo majhna v primerjavi z dinamiko višine tekočin v posodah in jo lahko zanemarimo (predpostavljamo, da ni zakasnitve med referenčno hitrostjo  $\omega_{1r}$  in pravo hitrostjo  $\omega_1$ ). V uporabljenem podsistemu imamo še dvopoložajni ventil V1 ter ročna ventila V2 in V3. Ventil V1 med posodama R1 in R2 je povsem odprt in omogoča pretok tekočine med dvema posodama, ventil V2 z zaprtjem loči izbrani podsistem od preostale naprave, ventil V3 pa je delno odprt, da omogoči odtekanje vode iz posode R2 v shranjevalnik R0. Presek shranjevalnika R0 je precej večji od presekov posod R1 in R2, zato je nivo tekočine v njem praktično konstanten.

Na izbran podsistem naprave vplivamo preko hitrosti vrtenja črpalke P1 in z odprtostjo ventila V2. Vhodi in izhodi naprave so povezani z osebnim računalnikom prek analogno/digitalnega (A/D) in digitalno/analognega (D/A) vmesnika *Burr-Brown PCI 20000 system*. Za vodenje naprave in opravljanje meritev je omogočena uporaba okolja Matlab/Simulink.

## Seznam senzorjev in aktuatorjev

Seznam merjenih spremenljivk izbranega podsistema naprave je viden v tabeli E.1. Nekaterne vrednosti izhodov senzorjev zaradi lažje preglednosti niso izražene z mednarodnim sistemom enot (SI). To velja za spremembe tlaka (cmH<sub>2</sub>O), velikosti pretokov (cm<sup>3</sup>/s) in nivoje tekočine v posodah (cm).

**Tabela E.1:** Tabela spremenljivk modela

Simbol	Spremenljivka	Enota	Področje	Opis
LT1	$h_1$	cm	0 - 60	nivo tekočine v posodi R1
LT2	$h_2$	cm	0 - 60	nivo tekočine v posodi R2
DPT1	$p_1$	cmH <sub>2</sub> O	0 - 100	tlačna razlika na črpalki P1
FT1	$Q_1$	cm <sup>3</sup> /s	0 - 30	pretok skozi črpalko P1
-	$\omega_1$	-	0 - 6	“hitrost” vrtenja črpalke P1
-	$U_1$	V	0 - 10	napetost na motorju črpalke P1
-	$I_1$	-	0 - 0.3	tok skozi motor črpalke P1
-	$s_1$	-	0,10	položaj dvopoložajnega ventila V1

Ker zaradi konstrukcije črpalke P1 ni mogoče pridobiti podatka o njeni hitrosti  $\omega_1$  neposredno, jo merimo 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 prek frekvenčno-napetostnega pretvornika v napetost na izhodu senzorja. Referenco za regulator kotne hitrosti  $\omega_{1r}$  črpalke P1 določamo z napetostjo v območju 0 – 6 V. Črpalka vsebuje povratnozančno regulacijo, ki je dovolj hitra, da lahko razliko med referenčno in pravo hitrostjo motorja, ki jo poganja, v primerjavi z dinamiko spreminjanja višine tekočine v posodah obravnavamo kot zanemarljivo. Meritve nivojev se opravljajo s kapacitivnimi senzorji, ki imajo linearno karakteristiko. Meritev pretokov ne uporabljamo.

Na izhodu A/D vmesnika so vse veličine predstavljene z napetostnim signalom. Zanimali nas bosta predvsem spremenljivki, ki ju bomo uporabljali pri identifikaciji naprave: nivo tekočine  $h_2$  v posodi R2, pretvorjen v [cm], in hitrost motorja črpalke P1, izražena z napetostjo v [V].

### Matematični model podsistema naprave

Za boljše razumevanje procesa bo ta predstavljen še z matematičnim modelom. Razvoj matematičnega modela naprave je podrobno opisan v [19].

Del naprave, ki nas zanima, lahko predstavimo z dvema stanjema v prostoru stanj: z nivojem tekočine  $h_1$  v posodi R1 in nivojem tekočine  $h_2$  v posodi R2. Nivo  $h_2$  je hkrati tudi izhod sistema, vhod pa napetost  $U_1$  na motorju črpalke P1, ki uravnava njeno hitrost  $\omega_1$ . V matematični obliki lahko sistem predstavimo:

$$\begin{aligned}
 \dot{h}_1 &= \frac{1}{A_1} \left( Q_1(\omega_1, h_1) - Q_{12}(h_1, h_2) \right), \\
 \dot{h}_2 &= \frac{1}{A_2} \left( Q_{12}(h_1, h_2) - Q_2(h_2) \right), \\
 Q_1(h_1, \omega_1) &= k_{1F}(h_{1\text{stat}}(\omega_1) - h_1) + k_{2F} \sqrt{h_{1\text{stat}}(\omega_1) - h_1}, \\
 Q_{12}(h_1, h_2) &= k_{V1} \sqrt{\rho g (h_1 - h_2)}, \\
 Q_2(h_2) &= k_{V3} \sqrt{\rho g h_2}, \\
 k_{1F}(\omega_1) &= k_{1F0} + k_{1F1}\omega_1 + k_{1F2}\omega_1^2, \\
 k_{2F}(\omega_1) &= k_{2F0} + k_{2F1}\omega_1,
 \end{aligned} \tag{E.1}$$

kjer sta  $A_1$  in  $A_2$  preseka posod R1 in R2,  $k_{V1}$  konstanta ventila V1,  $k_{V3}$  konstanta ventila V3,  $k_{1F}$  in  $k_{2F}$  parametra črpalke P1,  $h_{1\text{stat}}(\omega_1)$  pa ustaljena višina  $h_1$  pri neki hitrosti vrtenja črpalke  $\omega_1$  iz statične karakteristike  $(\omega_1, h_{1\text{stat}}(\omega_1))$ . Vrednosti konstantnih parametrov so v tabeli E.2.

Enačbe (E.1), ki opisujejo identificirani podsistem naprave, vsebujejo naslednje nelinearnosti: korenske odvisnosti tokov tekočine  $Q_1, Q_2$  in  $Q_3$  v odvisnosti od

višin tekočine v obeh posodah  $h_1$  in  $h_2$  ter nelinearno odvisnost vhodnega toka tekočine  $Q_1$  od hitrosti motorja v črpalki  $\omega_1$  in ustaljene višine tekočine pri tej hitrosti  $h_{1\text{stat}}(\omega_1)$ .

Opisani matematični model je bil narejen za analizo naprave, zato v njem niso vključene prav vse nelinearnosti. Poudarili bi predvsem dve neopisani nelinearnosti: karakteristiko črpalke in karakteristike kapacitivnih senzorjev za merjenje nivoja tekočine.

**Tabela E.2:** Tabela konstantnih parametrov modela

Konstanta	Vrednost	Enota	Opis
$A_1$	19.16	cm <sup>2</sup>	preseki posode R1
$A_2$	19.16	cm <sup>2</sup>	preseki posode R2
$k_{1F0}$	0.6595		koeficient v izrazu za $k_{1F}$
$k_{1F1}$	-0.2128		koeficient v izrazu za $k_{1F}$
$k_{1F2}$	$1.751 \cdot 10^{-2}$		koeficient v izrazu za $k_{1F}$
$k_{2F0}$	0.9352		koeficient v izrazu za $k_{2F}$
$k_{2F1}$	0.3024		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
$k_{V3}$	ni pod.		tokovni koeficient v smeri od R2 do R0

## E.2 Proces priprave plina

Proces priprave plina je ena izmed enot poliindustrijskega procesnega laboratorija Odseka za sisteme in vodenje na Institutu Jožef Stefan [61, 62, 114].

Procesni laboratorij je nastal ob podpori evropskega programa TEMPUS ALIAC (angl. *Active Learning In Automatic Control*) in je opremljen z industrijsko procesno opremo in napravami industrijskih razsežnosti. Predstavlja izvor različnih inženirskih nalog in problemov v zvezi z avtomatskim vodenjem procesov ter poligon za praktično preizkušanje različnih metod s področja procesnega vodenja. Dograjuje se glede na trenutne potrebe, pri čemer se uporablja komercialno dostopna profesionalna industrijska procesna oprema, ki omogoča razvoj in preizkus delovanja sodobnih metod avtomatskega vodenja ob upoštevanju omejitev, neidealnosti in tehničnih posebnosti, na katere naletimo v industriji.

Laboratorij je sestavljen iz dveh tehnoloških sklopov – procesnih enot, ki predstavljata dva tipična industrijska procesa. Delujeta lahko samostojno ali medsebojno povezano, saj je omogočen pretok materiala, energije in informacije med njima. Nas zanima samo proces priprave plina, slika E.2, za katerega je realiziran vmesnik za vodenje v realnem času iz okolja Matlab/Simulink. Proces kot samostojni del nima pomembnejše funkcije, v povezavi z drugimi deli procesnega laboratorija je namenjen:

- zajemanju in ohlajanju dimnih plinov — pri tem se dimni plini v prirejenem injektorju mešajo s hladilno vodo in
- ločevanju mešanice hladilne vode in dimnih plinov na ponovno uporabljivo vodo in na dimne pline pod tlakom, primernim za proces kemijske nevtralizacije bazičnih odplak.

Dimni plini morajo biti na izhodu naprave ohlajeni in pod ustreznim tlakom – približno 0.5 bara.

### Procesna oprema enote za pripravo plina

Procesna enota za pripravo plina vsebuje naslednjo procesno opremo, slika E.3:

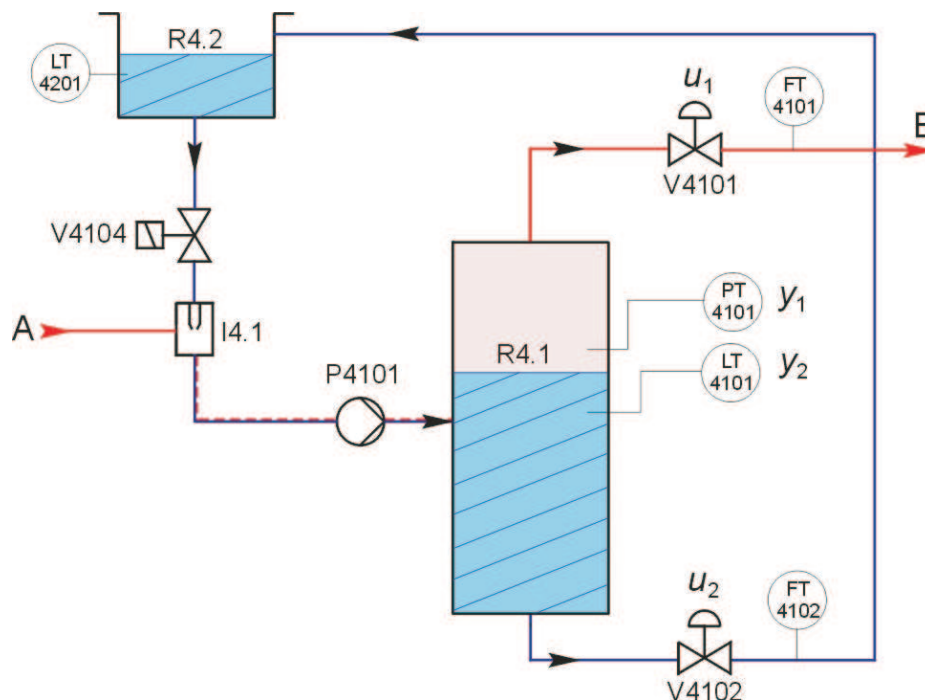
- ločevalnik plina in tekočine (krajše ločevalnik) R4.1, ki je opremljen z:
  - zveznim ventilom za plin V4101,



**Slika E.2:** Fotografija procesa priprave plina

- merilnikom pretoka plina iz ločevalnika FT4101,
- zveznim ventilom za vodo V4102,
- merilnikom pretoka vode iz ločevalnika FT4102,
- analognim merilnikom nivoja vode LT4101,
- merilnikom tlaka plina v ločevalniku PT4101,

- shranjevalna posoda R4.2 je opremljena z:
  - analognim merilnikom nivoja vode LT4201,
  - nivojskim stikalom za indikacijo maksimalnega nivoja LS4201,
- dvopoložajni ventil V4104,
- injektor I4.1,
- črpalka na vodni obroč P4101 s frekvenčnim regulatorjem.



Slika E.3: Shema procesa priprave plina

### Delovanje enote za pripravo plina

Priprava plina temelji na odvzemu dimnih plinov iz dimnika peči s frekvenčno regulirano črpalko P4101. Črpalka črpa vodo iz shranjevalne posode R4.2 v tlačno posodo R4.1. Voda v injektorju iz vhoda A v cevovod povleče dimne pline, ta mešanica potuje naprej do tlačne posode, kjer se voda in plin ločita. Zaradi prisiljenega dotoka mešanice vode in plina v posodo in regulirnih ventilov V4101 in V4102 lahko v tlačni posodi dosežemo nadtlak, ki vodo skozi izpust na dnu ločevalnika in ventil V4102 potisne nazaj v zbiralno posodo, s čimer je tokokrog vode sklenjen. Na izhodu B iz ločevalnika prek ventila V4101 v nadaljnji

proces kemijske nevtralizacije pod tlakom izstopa plin. Tlak plina v zgornjem delu ločevalnika mora biti večji od hidrostatičnega tlaka med nivojem vode v shranjevalni posodi in nivojem vode v ločevalniku, da je omogočen odtok vode iz ločevalnika prek regulacijskega ventila V4102 v shranjevalno posodo. Nivo hladilne vode v ločevalniku reguliramo z zveznim ventilom V4102, tlak plina pa z zveznim ventilom V4101 ali s frekvenčno regulirano črpalko P4101. V shranjevalni posodi merimo nivo vode z analognim merilnikom nivoja LT4201. Po potrebi dotočimo vodo prek ročnega ventila, morebitno odvečno vodo pa zaznavamo z nivojskim stikalom LS4201 in jo prek preliva pretočimo v kanalizacijo. Če hladilna voda zelo dolgo kroži po zaključeni poti v sistemu, obstaja nevarnost pregretja vode [102, 114].

### **Povezava enote za pripravo plina z vmesnikom za eksperimentiranje**

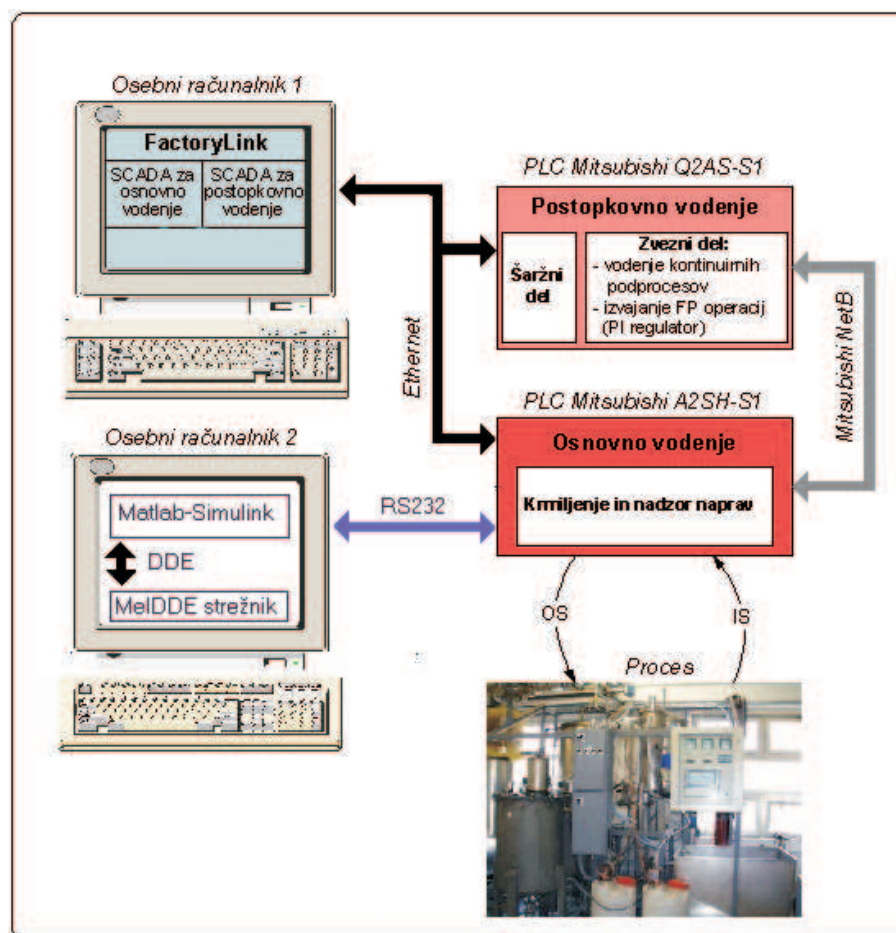
Tu bosta na kratko predstavljena koncept vodenja procesa in izvedba komunikacije med procesno enoto za pripravo plina in vmesnikom za eksperimentiranje. Povzeta sta iz [61], kjer sta tudi podrobneje predstavljena.

Koncept vodenja procesa in izvedba komunikacije sta predstavljena na sliki E.4. Sistem vodenja procesa priprave plina poleg procesa sestavljajo naslednji deli:

- senzorji in aktuatorji za zajem podatkov in vplivanje na proces;
- dva programirljiva logična krmilnika (angl. *Programmable Logic Controller*, PLC) podjetja Mitsubishi:
  - Mitsubishi A2SH-S1 za osnovno vodenje;
  - Mitsubishi Q2AS-S1 za postopkovno vodenje;
- dva osebna računalnika:
  - eden je namenjen SCADA nadzornemu sistemu;
  - drugi pa eksperimentiranju s procesom iz okolja Matlab/Simulink.

Procesno manj zmogljiv PLC Mitsubishi A2SH-S1 je namenjen osnovnemu vodenju. Vsebuje algoritme osnovnega vodenja, s senzorji in aktuatorji v procesu je povezan prek analognih in digitalnih vhodov in izhodov. Procesno zmogljivejši PLC Mitsubishi Q2AS-S1 je namenjen postopkovnemu vodenju in algoritmom, ki operirajo z realnimi števili (npr. PI regulator). Programirljiva logična krmilnika sta med seboj povezana z Mitsubishijevo mrežo NetB.

Prvi osebni računalnik je namenjen implementaciji SCADA nadzornega sistema za osnovno in postopkovno vodenje, izdelanega s programskim paketom Factory



Slika E.4: Koncept vodenja procesa priprave plina [61].

Link. Sistem za nadzor in vodenje je v sistemu vodenja procesnega laboratorija namenjen zajemu procesnih podatkov in podatkov faz iz PLC-ja (stanja naprav in faz, procesne veličine) ter posredovanje podatkov procesnim napravam in fazam, npr. ukazov in parametrov. PLC-ja sta z nadzornim sistemom povezana prek Etherneta in ustreznih modulov, za komunikacijo pa skrbi vmesnik Mecom, ki je del Factory Linka.

Na drugem osebni računalniku sta nameščena paketa Matlab in MelDDE strežnik. Strežnik prek serijskega izhoda osebnega računalnika omogoča povezavo Matlab in krmilnika A2SH-S1 na podlagi DDE (angl. *Dynamic Data Exchange*) komunikacije, sistemsko podprte komunikacije v okolju Windows. Tako lahko na proces priprave plina vplivamo iz okolja Matlab/Simulink, pri pogoju, da je proces v stanju varnega delovanja; tega zagotavlja in nadzoruje krmilnik A2SH-S1.

Minimalni čas vzorčenja, ki ga s tako izvedbo lahko dosežemo, je odvisen od časa, potrebnega za DDE komunikacijo, in časa, potrebnega za simulacijo oz. potrebne izračune v Matlab/Simulinku; ta dva časa sta odvisna od zmogljivosti uporabljenega računalnika. Pri izbranem času vzorčenja  $T_s = 1$  s je izvedba vodenja delovala, čas vzorčenja pa bi po potrebi verjetno lahko zmanjšali.

### Matematični opis procesa v ločevalniku plina in tekočine

Za prikaz pogloblitnih relacij v procesu in za ilustracijo nelinearnosti procesa lahko proces priprave plina opišemo z dvema enačbama:

$$\begin{aligned} \frac{dp_1}{dt} &= \frac{1}{S_1(h_{T_1} - h_1)} (p_0(\alpha_0 + \alpha_1 p_1 + \alpha_2 p_1^2 - k_1 R_1^{u_1-1} \sqrt{p_1}) \\ &\quad + (p_0 + p_1)(\Phi_w - k_2 R_2^{u_2-1} \sqrt{p_1 + k_w(h_1 - h_{T_2})})), \\ \frac{dh_1}{dt} &= \frac{1}{S_1} (\Phi_w - k_2 R_2^{u_2-1} \sqrt{p_1 + k_w(h_1 - h_{T_2})}), \end{aligned} \tag{E.2}$$

kjer je  $u_i$ ,  $i = 1, 2$ , vhodni signal za ventil V410*i*;  $h_i$ ,  $i = 1, 2$ , višina tekočine v posodi R4.*i*;  $p_1$  relativni pritisk zraka v posodi R4.1;  $S_i$ ,  $i = 1, 2$ , presek posode R4.*i*;  $p_0$  zračni pritisk;  $h_{T_i}$ ,  $i = 1, 2$ , višina posode R4.*i*;  $R_i$ ,  $i = 1, 2$ , razmerje pretoka med zaprtim in odprtim ventilom V410*i*;  $k_i$ ,  $i = 1, 2$ , koeficient pretoka ventila V410*i*;  $\Phi_w$  znana konstanta vodnega toka skozi črpalko in  $\alpha_i$ ,  $i = 1, 2, 3$  konstante.

Iz enačb lahko sklepamo, da je opisovani nelinearni proces multivariabilen z dvema vhodoma, dvema izhodoma in navzkrižnimi povezavami. Če imamo povratnozančno regulacijo nivoja tekočine  $h_1$ , lahko ta sistem predstavimo kot univariabilen sistem z vhomom  $u_1$  in izhodom  $p_1$ . Iz enačb (E.2) lahko tudi razberemo, da je tlak  $p_1$  nelinearno odvisen od nivoja  $h_1$  in vhodnega toka, zaradi česar se proces v različnih območjih obnaša različno.

## Dodatek F

# Zasnova programskega modula za identifikacijo z GP modeli

V tem dodatku je predstavljena zasnova programskega modula v okolju Matlab za uporabo Gaussovih procesov pri identifikaciji dinamičnih sistemov, s čimer želimo olajšati delo potencialnim uporabnikom metode v identifikaciji.

Pri pisanju zasnove programskega paketa smo si pomagali s programskim paketom *GPML*, ki sta ga napisala C. E. Rasmussen in C. K. I. Williams in je za akademske potrebe prosto dostopen<sup>1</sup>, ter delom, ki so ga pred tem opravili C. E. Rasmussen, R. Murray-Smith in J. Kocijan.

Na internetni strani <http://www.gaussianprocess.org/> so na voljo tudi nekateri drugi programski paketi v različnih jezikih, vendar noben od teh ni prirejen za uporabo pri identifikaciji in simulaciji dinamičnih sistemov.

Predstavili bomo seznam funkcij modula, ki bodo razdeljene po sklopih delovanja, in potrebni Matlab/Simulink model. Prispevek predlaganega osnutka je predvsem v predlogih funkcij, ki naj bi bile uporabljane za simulacijo dinamičnih sistemov. Seznam funkcij ni dokončen in ga je mogoče razširiti predvsem z dodatnimi kovariančnimi funkcijami<sup>1</sup> in z dodatnimi rutinami za hitrejše učenje GP modelov, omenjenimi v drugem poglavju disertacije.

---

<sup>1</sup>GPML programski paket, <http://www.gaussianprocess.org/gpml>, pod pogoji uporabe, zapisanimi v datoteki.

## Pregled potrebnih Matlabovih funkcij in modelov za uporabo GP modelov za identifikacijo dinamičnih sistemov

Osnovne funkcije za uporabo GP modela	
<code>gpr</code>	- napovedovanje GP modela (statični sistemi) ter - vračanje verjetnosti učnih podatkov in njenih odvodov po posameznih hiperparametrih
<code>traingp</code>	učenje GP modela
<code>minimize</code>	iskanje lokalnega optimuma diferenciable funkcije več spremenljivk, uporablja jo funkcija <code>traingp</code>

Kovariančne funkcije	
<code>covConst</code>	konstantna kovariančna funkcija
<code>covLINard</code>	linearna kovariančna funkcija z ARD lastnostjo
<code>covLINone</code>	linearna kovariančna funkcija z enotnim hiperparametrom
<code>covSEard</code>	Gaussova kovariančna funkcija z ARD lastnostjo (2.4)
<code>covSEiso</code>	Gaussova kovariančna funkcija z izotropično mero razdalje
<code>covPeriodic</code>	periodična kovariančna funkcija
<code>covNoiseWht</code>	kovariančna funkcija, ki izraža prispevek belega šuma
<code>covNoiseClr</code>	kovariančna funkcija, ki izraža prispevek barvnega šuma
<code>covSum</code>	seštevanje kovariančnih funkcij
<code>covProd</code>	množenje kovariančnih funkcij

Simulacija GP modela	
<code>simgpnaive</code>	simulacija GP modela brez propagacije negotovosti
<code>simgpexact</code>	simulacija GP modela s propagacijo negotovosti z anal. aproksimacijo
<code>simgpnumer</code>	simulacija GP modela s propagacijo negotovosti z numer. aproksimacijo

Pomožne funkcije	
<code>premmx</code>	normiranje podatkov, da ti ležijo med -1 in 1
<code>postmmx</code>	“odnormiranje” normiranih podatkov
<code>makesiminput</code>	konstrukcija matrike vhodnih vektorjev za simulacijo modela
<code>plotgp</code>	grafični prikaz rezultatov simulacije (primerjava odziva modela in zelenega odziva skupaj s 95% pasom zaupanja)
<code>plotgpe</code>	grafični prikaz napake skupaj s 95% pasom zaupanja
<code>loss</code>	kvantitativna ocena rezultatov simulacije (rezultati cenilk)

<b>Funkcije za uporabo z LMGP modelom</b>	
<code>lmgp</code>	- napovedovanje LMGP modela (statični sistemi) ter - vračanje verjetnosti učnih podatkov in njenih odvodov po posameznih hiperparametrih
<code>trainlmgp</code>	učenje LMGP modela
<code>simlmgpnaive</code>	simulacija LMGP modela brez propagacije negotovosti
<code>simlmgpexact</code>	simulacija LMGP modela s propagacijo negotovosti z analitično aproksimacijo
<code>simlmgpnumer</code>	simulacija LMGP modela s propagacijo variance z numerično aproksimacijo

<b>Funkcije za uporabo s FSGP modelom</b>	
<code>trainfsgp</code>	učenje FSGP modela
<code>fsgpmodelfun</code>	napovedovanje izhoda FSGP modela za uporabo v Simulinku
<code>fsgpcontrolfun</code>	predloga funkcije za napovedovanje regularne veličine
<code>getfsgpparams</code>	napovedovanje parametrov FSGP modela

<b>Matlab/Simulink model za uporabo s FSGP modelom</b>	
<code>fsgpmodel</code>	Matlab/Simulink model, ki je predloga za simulacijo, načrtovanje vodenja in vodenje FSGP modela



# Izjava

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

V Ljubljani, 17.9.2007.

Kristjan Ažman