Automatic Model Selection Algorithm Based on BYY Harmon Learning for Mi ture of Gaussian Process Functional Regressions Models

Xia g a g G , Ta H g, a d Ji e Ma(✉)

De a tme t f I f mati a d C m tati al Scie ce , Sch l f Mathematical Scie ce a d LMAM, Peki g U i e it , Beiji g 100871, Chi a

Abstract. F ite mixt e m del , dete mi i g the mbe f c m e t i efe ed t a m del electi . Thi a e t f a d a a t matic m del elec- ti alg ithm ba ed Ba e ia Yi g-Ya g (BYY) ha m lea i g f mixt e f Ga ia ce f cti al eg e i (mix-GPFR) m del . BYY ha m lea i g ha bee cce f ll a lied t the m del electi blem f Ga - ia mixt e m del (GMM), b t it ca t be di ectl ed f that f mix-GPFR m del . We d t the ca e f thi blem a d e ac i g mecha i m f c e ec t cti ba ed Ga ia ce (GP) m del , th gh hich, e t a f m a mix-GPFR m del i t a GMM. Th , e ca make m del electi f mix-GPFR m del ia BYY ha m lea i g. E e ime tal e lt h that alg ithm ca d the timal mbe f c m e t i a m lti- ce c e data et.

Ke ords: Mixt e f Ga ia P ce F cti al Reg e i · M del Selecti · Ba e ia Yi g-Ya g Ha m Lea i g · C e Rec t cti

1 Introduction

Ga ia ce (GP) m del a e a effecti e t l f Ba e ia li ea a amet- ic cla i cati a d eg e i , e.g., cla if i g the image f ha d itte digit a d m deli g the i e e d amic f a b t a m [1]. H e e , the ca t deal ith m lti- ce c e data et effecti el . T e c me thi limitati , mixt e f Ga ia - ce f cti al eg e i (mix-GPFR) m del e e ed [2, 3] a d the e te i e e ea ch ha bee de ted t e timati g thei a amete , a al i g thei e f ma ce, a d a l i g them t eal- ld blem [4 8].

Like the ite mixt e m del , mix-GPFR m del al face the blem f m del electi , amel dete mi i g the mbe f Ga ia ce f cti al eg e - i (GPFR) c m e t . Si ce a i a iate mbe f GPFR c m e t ill i e itabl lead t ge e ali ati abilit , m del electi i f g eat im ta ce. I additi t maki g m del electi tili i g d mai k ledge e e ie ce, e ca

D.-S. H a g et al. (Ed .): ICIC 2023, LNAI 14089, . 391 403, 2023.
htt ://d i. g/10.1007/978-981-99-4752-2_33

al de ig a t matic m del electi alg ithm . The t aditi al meth d i t ch e the timal mbe f GPFR c m e t th gh ce tai tati tical electi c ite i . F ex. am le, Qia g et al. [6] ed the liti g ex. ectati -max. imi ati (SEM) alg ithm ba ed the Ba e ia i f mati c ite i (BIC) [9]. H e e , all the ex. i t- i g tati tical electi c ite ia fte ca e a im e mbe f GPFR c m e t a d the e f a tati tical electi c ite i ic a high time c m lex. it , i ce e eed t e eat the h le a amete e timati g ce f diffe e t mbe f GPFR c m e t . M e e , t cha tic im lati meth d , ch a e e ible j m Mar k chai M te Ca l [10] a d Di ichlet ce e [11], ha e al bee ed t deal ith the m del electi b lem f mix.-GPFR m del [5, 7, 8]. H e e , the e meth d e ic llecti g a la ge mbe f am le , hich e lt i a high c m tati al c t.

F Ga ia mix.t e m del (GMM), the a t matic m del electi alg ithm ba ed Ba e ia Yi g-Ya g (BYY) ha m lea i g [12, 13] ha e ac ie ed bette e lt a d highe c m tati eed tha th e ba ed tati tical electi c ite ia a d t cha tic im lati meth d [14 20

He e, i ce the i dicat c ariable i cala , e de te it a z i tead f . F the GMM, e e tabli the f ll i g BYY tem: $q(z=g)=\pi_g$; $q(\ |z=g)=\mathcal{N}(\ |\boldsymbol{\mu}_g,\boldsymbol{\Sigma}_g)$; $p(\ )=\frac{1}{I}\sum_{i=1}^{I}\delta(\ -\ _i)$, i.e. the em ical de it f cti ;

$$p(z=g|\ )=\frac{\pi_g\mathcal{N}(\ |\boldsymbol{\mu}_g,\boldsymbol{\Sigma}_g)}{\sum_{s=1}^{G}\pi_s\mathcal{N}(\ |\boldsymbol{\mu}_s,\boldsymbol{\Sigma}_s)}. \tag{3}$$

M ce e , e ig ce the eg lari ati tem r, i.e. et $r=1$. The , e ha e

$$H(p||q)=J(\Theta)=\frac{1}{I}\sum_{i=1}^{I}\sum_{g=1}^{G}\frac{\pi_g\mathcal{N}(\ _i|\boldsymbol{\mu}_g,\boldsymbol{\Sigma}_g)}{\sum_{s=1}^{G}\pi_s\mathcal{N}(\ _i|\boldsymbol{\mu}_s,\boldsymbol{\Sigma}_s)}1\ \big(\pi_g\mathcal{N}(\ _i|\boldsymbol{\mu}_g,\boldsymbol{\Sigma}_g)\big), \tag{4}$$

he e $J(\Theta)$ i called ham f cti a d $\Theta=\{\pi_g,\boldsymbol{\mu}_g,\boldsymbol{\Sigma}_g\}_{g=1}^{G}$.

Acc di g t BYY ham lea i g, the maximum f $J(\Theta)$ c e d t the timal mbe f Ga ia c m e t a d the be t a ameter [14 20]. He ce, e ca make m del electi a d e timate the a ameter b maximi i g $J(\Theta)$. I the ce f maximi i g $J(\Theta)$, the mix i g c cti f the ed da t Ga ia c m e t c e ge t e . C m ared ith the a t matic m del electi alg ithm ba ed tati ical electi c iteria a d t cha tic im lati meth d , th e ba ed BYY ham lea i g ha e ac ie ed bette ce lt a d highe c m tati eed [14 20].

3 Automatic Model Selection Algorithm Based on BYY Harmon Learning

Fi tl , e b ie i t d ce the mix.-GPFR m del. A GP i a c llecti f a d m a iable , a ite b et f hich i bject t a Ga ia di ib ti [1]. T ecif a GP $\{f(\ )|\ \in\mathcal{X}\subseteq\mathbb{R}^D\}$, e l eed t dete mi e it mea f cti $m(\ )$ a d c a ia ce f cti $c(\ ,\ ')$, he e.

$$m(\ )=\mathbb{E}[f(\ )]\ a d\ c(\ ,\ ')=\mathbb{E}[(f(\ )-m(\ ))(f(\ ')-m(\ '))]. \tag{5}$$

he e the GP i de ted a

$$f(\ )\sim\mathcal{GP}(m(\ ),c(\ ,\ ')). \tag{6}$$

I mix.-GPFR m del , ice $D=1$, e de te the i t a x i tead f . The , a mix.-GPFR m del ith G GPFR c m e t ca be e tabli hed th gh the f ll i g f m lae:

$$q(z=g)=\pi_g,\ where\ \pi_g\geq 0\ a d\ \sum_{g=1}^{G}\pi_g=1; \tag{7}$$

$$q(y(x)|z=g)=\mathcal{GPFR}(x|\mathbf{b}_g,\boldsymbol{\theta}_g,r_g)=\mathcal{GP}\Big(m(x|\mathbf{b}_g),c(x,x'|\boldsymbol{\theta}_g)+r_g^{-1}\delta(x,x')\Big). \tag{8}$$

I E . (8), $\delta(x, x')$ i the K: ecke: delta f cti ,

$$m(x|\mathbf{b}_g) = \varphi(x)^T \mathbf{b}_g \text{ a d } c(x, x'|\boldsymbol{\theta}_g) = \theta_{g0}^2 \text{ex} \left\{ -\frac{(x - x')^2}{2\theta_{g1}^2} \right\}, \tag{9}$$

here $\varphi(x) = [\varphi_1(x), \varphi_2(x), \ldots, \varphi_P(x)]^T$ i a c l m ect : f B- li e [21] a d $c(x, x'|\boldsymbol{\theta}_g)$ i : efe: : ed t a the . a: ed ex. e tial c a: ia ce f cti . θ_{g0}, θ_{g1}, a d r_g a: e iti e a: amete: .

The Yi g machi e f the mix. -GPFR m del i

$$q(z = g, y(x)) = q(z = g)q(y(x)|z = g) = \pi_g \mathcal{GPFR}(x|\mathbf{b}_g, \boldsymbol{\theta}_g, r_g) \tag{10}$$

a d it Ya g machi e i

$$p(z = g, y(x)) = p(y(x))p(z = g|y(x)) = p(y(x)) \frac{\pi_g \mathcal{GPFR}(x|\mathbf{b}_g, \boldsymbol{\theta}_g, r_g)}{\sum_{s=1}^{G} \pi_s \mathcal{GPFR}(x|\mathbf{b}_s, \boldsymbol{\theta}_s, r_s)}. \tag{11}$$

We de te a t: ai i g c : e data et a $\mathcal{D} = \{\mathcal{C}_i\}_{i=1}^{I}$, he: e $\mathcal{C}_i = \{(x_{in}, y_{in})\}_{n=1}^{N_i}$: e: e e t a t: ai i g c : e f le gth N_i. It i ge e: all a med that $x_{i1}, \ldots, x_{iN_i}$ a: e : a d m l di t: ib ted i the i te: al $[x_{\min}, x_{\max}]$ $(i = 1, \ldots, I)$. Let $_i = [x_{i1}, \ldots, x_{iN_i}]^T$, $_i = [y_{i1}, \ldots, y_{iN_i}]^T$, a d $\Theta = \{\pi_g, \mathbf{b}_g, \boldsymbol{\theta}_g, r_g\}_{g=1}^{G}$. F : the mix. -GPFR m del,

$$H(p||q) = \sum_{g=1}^{G} \int p(y(x))p(z = g|y(x)) \, (q(z = g)q(y(x)|z = g))\text{dy}(x) \tag{12}$$

ca t be a : x. imated b

$$J(\Theta) = \frac{1}{I} \sum_{i=1}^{I} \sum_{g=1}^{G} \frac{\pi_g \mathcal{N}(\ _i|\mathbf{m}_{ig}, \mathbf{C}_{ig})}{\sum_{s=1}^{G} \pi_s \mathcal{N}(\ _i|\mathbf{m}_{is}, \mathbf{C}_{is})} \, (\pi_g \mathcal{N}(\ _i|\mathbf{m}_{ig}, \mathbf{C}_{ig})) \tag{13}$$

ith $\mathbf{m}_{ig} = m(\ _i|\mathbf{b}_g)$ a d $\mathbf{C}_{ig} = c(\ _i, \ _i|\boldsymbol{\theta}_g) + r_g^{-1}\mathbf{I}_{N_i}$, he: e $\mathbf{I}_{N_i}$ i the: te[illegible]TD .1666Tc (,...,)Tj /F7

$\hat{\mathcal{C}}_i = \{(x_n, \hat{y}_{in})\}_{n=1}^N$ fr m $\hat{f}_i(x)$ ith $x_n = x_{\mathrm{mi}} + (n-1)\Delta$. D ri g am li g, the ar ia ce f Ga ia i e i

$$\sigma_1^2 = \frac{1}{N_i}\sum_{n=1}^{N_i}\left(y_{in} - \hat{f}_i(x_{in})\right)^2. \tag{14}$$

It i clear that

$$\sigma_2^2 = \frac{1}{N_i}\sum_{n=1}^{N_i}(y_{in} - f_i(x_{in}))^2 \tag{15}$$

i a bia ed e timate f the ar ia ce f Ga ia i e i the r ce f am li g $\mathcal{C}_i$ fr m $f_i(x)$. He ce, σ_1^2 i a g d e timate f the ar ia ce the a m ti that there are ig i ca t differe ce bet ee $f_i(x)$ a d $\hat{f}_i(x)$. I t iti el , $\hat{\mathcal{C}}_i$ i a g d a r ximati f $\mathcal{C}_i$. That i t a , the differe ce bet ee the teri r mea f cti r ec er ed fr m $\hat{\mathcal{C}}_i$ a d $\mathcal{C}_i$, re ecti el , i mall, hich ill be alidated thr gh ex erime t i Sect. 4.

Let $\hat{\mathcal{D}} = \{\hat{\mathcal{C}}_i\}_{i=1}^{I}$, $= [x_1, x_2, \ldots, x_N]^T$, a d $\hat{}_i = [\hat{y}_{i1}, \hat{y}_{i2}, \ldots, \hat{y}_{iN}]^T$. $\hat{}_i$ ca be regarded a a am le f the f ll i g GMM:

$$q(z=g) = \pi_g \text{ here } \pi_g \geq 0 \text{ a d } \sum_{g=1}^{G}\pi_g = 1;\; q(\hat{}|z=g) = \mathcal{N}(\hat{}|\mathbf{m}_g, \mathbf{C}_g), \tag{16}$$

here $\mathbf{m}_g = m(|\mathbf{b}_g)$ a d $\mathbf{C}_g = c(, |\boldsymbol{\theta}_g) + r_g^{-1}\mathbf{I}_N$. It Yi g machi e i

$$q(z=g, \hat{}) = q(z=g)q(\hat{}|z=g) = \pi_g\mathcal{N}(\hat{}|\mathbf{m}_g, \mathbf{C}_g) \tag{17}$$

a d it Ya g machi e i

$$p(z=g, \hat{}) = p(\hat{})p(z=g|\hat{}) = p(\hat{})\frac{\pi_g\mathcal{N}(\hat{}|\mathbf{m}_g, \mathbf{C}_g)}{\sum_{s=1}^{G}\pi_s\mathcal{N}(\hat{}|\mathbf{m}_s, \mathbf{C}_s)}. \tag{18}$$

The , it c rre di g harm f cti i

$$J(\Theta) = \frac{1}{I}\sum_{i=1}^{I}\sum_{g=1}^{G}\frac{\pi_g\mathcal{N}(\hat{}_i|\mathbf{m}_g, \mathbf{C}_g)}{\sum_{s=1}^{G}\pi_s\mathcal{N}(\hat{}_i|\mathbf{m}_s, \mathbf{C}_s)} 1\ \left(\pi_g\mathcal{N}(\hat{}_i|\mathbf{m}_g, \mathbf{C}_g)\right). \tag{19}$$

A i the ca e ith GMM , the max im m f $J(\Theta)$ c rre d t the timal mber f GPFR c m e t a d the be t arameter . Theref re, e ca make m del electi a d lear the arameter b maximi i g $J(\Theta)$ thr gh merical timi ati meth d .

After the trai i g r ce , e ca determi e the cla f a trai i g c r e acc rdi g t the maxim m a teri ri r babilit , i.e. let

$$z_i = \arg\max_{g\in\{1,2,\ldots,G\}} \frac{\pi_g\mathcal{N}(\hat{}_i|\mathbf{m}_g, \mathbf{C}_g)}{\sum_{s=1}^{G}\pi_s\mathcal{N}(\hat{}_i|\mathbf{m}_s, \mathbf{C}_s)} (i = 1, 2, \ldots, I). \tag{20}$$

The r ed da t GPFR c m e t d t get a trai i g c r e d e t their er mall mixi g r rti . The cla f a te t c r e ca al be determi e i thi a . Be ide , f r a te t c r e, e ca redict the te t t t b calc lati g their c diti al di trib ti gi e the k t t . The detail are referred t [4 8].

4 E perimental Results

I thi ecti , e e i e thetic data et a d real- ld data et t e if the effecti e e f c ed a t matic m del electi alg ithm. We c m a e mix.-GPFR m del t ai ed ia c ed alg ithm ith GP m del , mix.-GP m del , GPFR m del , a d mix.-GPFR m del t ai ed th gh the t aditi al EM alg ithm [2, 3] a d the SEM alg ithm [6].

Si ce e a e mai l c ce ed ith the edicti abilit f mix.-GPFR m del , the t ted mea a e e (RMSE) i ch e a the e al ati metic. It i a med that the e a e T te t c e a d the te t t t f the tth ($t = 1, 2, \ldots, T$) te t c e a e $y_{t1}, y_{t2}, \ldots, y_{tM}$, h e c e di g edicti al e a e $\hat{y}_{t1}, \hat{y}_{t2}, \ldots, \hat{y}_{tM}$, e ecti el . It f ll that

$$\mathrm{RMSE} = \sqrt{\frac{1}{TM}\sum_{t=1}^{T}\sum_{m=1}^{M}\left(y_{tm} - \hat{y}_{tm}\right)^2}. \tag{21}$$

A a e tl , a malle RMSE i dicate a bette edicti e lt.

4.1 On S nthetic Datasets

The i e thetic data et a e de ted a $\mathcal{S}_2, \mathcal{S}_3, \ldots, \mathcal{S}_{10}$, e ecti el , he e the b- c i t ce e the mbe f m e t. F each m e t, e am le 20 t ai i g c e a d 10 te t c e f m a GP ith - e mea f cti . The mea f cti a d a amete f the Ga ia c ce e ed t ge e ate the i e thetic data et a e li t i Table 1, he e $\mathcal{S}_l$ ($l = 2, 3, \ldots, 10$) a e ge e ated b the t t l GP . Each c e c it f 100 i t , h ei t a e a d m l di t ib ted i $[-3, 3]$. The 60 i t the left ide f a te t c e a e k a d the 40 e the ight ide a e ed f te ti g.

Fi t l , e de m t ate the effecti e e f c e ec t cti ba ed GP m del th gh ex e i me t. A t ai i g c e i ca d m l ch e f m each m e t i $\mathcal{S}_9$. Fig e 1 e e t the ec t cti c e f the i et ai i g c e . Fig e 1 i m- ed f 9 b- g e , each f hich e e t a t ai i g c e, it ec t cti c e, a d thei te i mea f cti . A ca be ee f m the g e, alth gh the e a e ig i ca t diffe e ce bet ee a t ai i g c e a d it ec t cti c e, thei - te i mea f cti a e imila , hich im lie that c ed c e ec t cti ba ed GP m del i effecti e.

Whe te ti g c ed alg ithm, G i initiali ed a $l+3$ f $\mathcal{S}_l$. T ill t ate the bad effect f a g mbe f GPFR m e t edicti abilit , et ai mix.- GPFR m del c iti g f $l-1$ a d $l+1$ GPFR m e t ia the EM alg ithm [2, 3], hich a e de ted a mix.-GPFR (-1), a d mix.-GPFR (+1), e ecti el . Simila l , mix.-GP m del ith $l-1$ a d $l+1$ GP m e t a e de ted a mix.-GP (-1), a d mix.-GP (+1), e ecti el . Be ide , P i et t be 20. Table 2 e e t the e. e i me tal e. lt .

F m Table 2, e ee that the RMSE f the GPFR m del a d the mix.-GPFR m del a e malle tha th e f the GP m del a d the mix.-GP m del, e ecti el , hich

Table 1. Mea f cti a d a amete f the Ga ia ce e ed t ge e ate the i e thetic data et .

Mea f cti	$\boldsymbol{\theta}^T$	$\sqrt{r^{-1}}$
x^2	[0.5, 0.5]	0.15
$\left(-4(x+1.5)^2+9\right)1_{\{x<0\}}+\left(4(x-1.5)^2-9\right)1_{\{x\geq 0\}}$	[0.528, 0.4]	0.144
8 i $(1.5x-1)$	[0.556, 0.3]	0.139
i $(1.5x)+2x-5$	[0.583, 0.2]	0.133
i $(4x)-0.5x^2-2x$	[0.611, 0.1]	0.128
$-x^2$	[0.639, 0.1]	0.122
$\left(4(x+1.5)^2-9\right)1_{\{x<0\}}+\left(-4(x-1.5)^2+9\right)1_{\{x\geq 0\}}$	[0.667, 0.2]	0.117
5 c $(3x+2)$	[0.694, 0.3]	0.111
c $(1.5x)-2x+5$	[0.722, 0.4]	0.106
c $(4x)+0.5x^2+2x$	[0.75, 0.5]	0.1

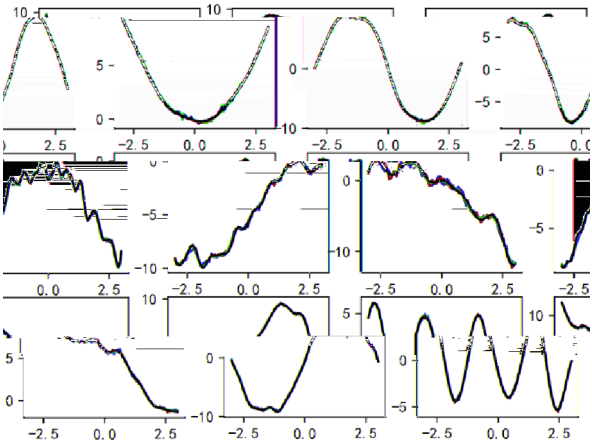

Fig. 1. The e lt f c e ec t cti the thetic data et . The ed, g ee , bl e, a d black c e e e e t the igi al c e, the ec t cted c e, the te i mea f cti f the igi al c e, a d the te i mea f cti f the ec t cted c e, e ecti el .

de m t ate the effecti e e f m deli g the mea f cti a a li ea c mbi ati f B- li e . B c m a i g the mix.-GP (mix.-GPFR) m del a d the GP (GPFR) m del, the eed f i t d ci g the mix.t e t ct e i de m t ated. F the m e, e ca ee that a g mbe f GPFR c m e t affect the edicti e lt badl . F $\mathcal{S}_2, \mathcal{S}_3, \ldots, \mathcal{S}_9$, b th the SEM alg ithm a d ed alg ithm d the c ect mbe f GPFR c m e t a d the i RMSE a e cl e. H e e, the time c m lex. it f the SEM alg ithm i highe tha that f ed alg ithm. O the e ha d,

the SEM algorithm need to decide the value of the parameter learning rate for different number of GPFR components. On the other hand, since different training data sets have different initialization, the computing time of the SEM algorithm has a high time complexity. For $\mathcal{S}_{10}$, since the SEM algorithm fail to find the true number of GPFR components, its RMSE is larger than that of the proposed algorithm.

Taking $\mathcal{S}_9$ for example, we present the clustering results of the proposed algorithm in Fig. 2, where different clusters are presented in different colors. On the left and right side of Fig. 2 are the clustering results of the proposed algorithm on the training and testing data sets, respectively. It is clear that the proposed algorithm correctly found all the components.

Table 2. RMSE and training time of all the methods on the synthetic data sets.

	$\mathcal{S}_2$		$\mathcal{S}_3$		$\mathcal{S}_4$	
	RMSE	Time (min)	RMSE	Time (min)	RMSE	Time (min)
GP	5.5831	6.87	4.7878	9.88	4.7798	13.09
mix.-GP (-1)	5.5239	6.12	4.6125	8.90	4.3580	15.84
mix.-GP (+1)	4.8240	7.39	4.6035	17.40	4.3488	23.31
GPFR	5.0759	6.68	4.6864	12.42	4.3051	17.72
mix.-GPFR (-1)	5.0214	8.07	0.9416	15.95	0.9510	18.93
mix.-GPFR (+1)	1.6846	14.20	1.0680	22.66	0.9319	25.79
mix.-GPFR (SEM)	**0.4312**	20.63	0.4856	41.59	0.5469	58.97
mix.-GPFR (BYY)	0.4401	9.46	**0.4746**	15.03	**0.5403**	18.64
	$\mathcal{S}_5$		$\mathcal{S}_6$		$\mathcal{S}_7$	
	RMSE	Time (min)	RMSE	Time (min)	RMSE	Time (min)
GP	4.9213	17.85	4.9897	15.07	5.3096	20.47
mix.-GP (-1)	4.4775	15.03	4.3834	20.14	4.3025	30.66
mix.-GP (+1)	4.5205	28.59	4.3813	29.19	4.3082	30.53
GPFR	4.8079	26.73	4.8649	24.25	4.9871	21.45
mix.-GPFR (-1)	0.8756	31.66	1.0776	29.67	1.3295	32.09
mix.-GPFR (+1)	0.9252	30.58	1.0270	35.37	1.0281	38.44
mix.-GPFR (SEM)	0.5638	81.34	**0.6057**	87.52	**0.6540**	92.78
mix.-GPFR (BYY)	**0.5573**	25.49	0.6137	23.66	0.6571	27.82
	$\mathcal{S}_8$		$\mathcal{S}_9$		$\mathcal{S}_{10}$	
	RMSE	Time (min)	RMSE	Time (min)	RMSE	Time (min)
GP	4.8180	19.67	4.4758	17.82	4.8438	21.67
mix.-GP (-1)	4.4904	24.13	4.1223	32.36	4.5730	32.78

(*continued*)

Table 2. (*continued*)

	$\mathcal{S}_2$		$\mathcal{S}_3$		$\mathcal{S}_4$	
	RMSE	Time (min)	RMSE	Time (min)	RMSE	Time (min)
mix.-GP (+1)	4.4818	23.70	4.1214	30.95	4.5878	28.14
GPFR	4.6871	26.73	4.3686	21.67	4.6865	20.97
mix.-GPFR (-1)	1.5325	33.78	1.0585	40.27	1.5279	41.56
mix.-GPFR (+1)	1.1891	35.96	0.9789	39.38	1.0343	49.78
mix.-GPFR (SEM)	0.6448	99.49	0.6233	116.85	1.4379	130.65
mix.-GPFR (BYY)	**0.6421**	28.91	**0.6199**	28.62	**0.6317**	32.46

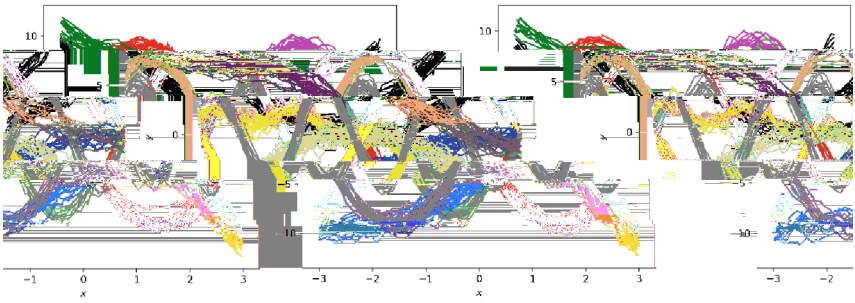

Fig. 2. Clustering results of automatic model selection algorithm on $\mathcal{S}_7$ and $\mathcal{S}_9$.

4.2 On Real-World Datasets

Here, we utilize the electricity load data set used by the Northeast Chinese Grid Company [8], which records electricity load every 15 minutes during 2009 and 2010. Hence, daily electricity load can be regarded as a curve with 96 dimensions. We divide the data set into two sub-data sets according to the year, which are referred to as $\mathcal{R}_1$ and $\mathcal{R}_2$, respectively. Each sub-data set consists of 200 training curves and 165 testing curves. Moreover, the 56 dimensions of the left side of the curves are used as input, and the 40 ones of the right side are used for testing.

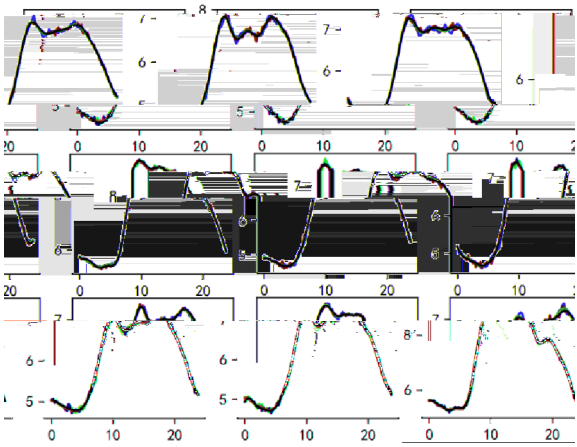

Fig. 3. The results of reconstructions on the electricity load data set. The red, green, blue, and black curves are the original curve, the reconstructed curve, the functional mean of the original curve, and the functional mean of the reconstructed curve, respectively.

Although all the curves have the same length, they treat them as if the data have the same length. Like the synthetic data set, we add small changes to training curves in $\mathcal{R}_1$, which curves are denoted in Fig. 3. A can be seen from the figure, curves denoted curves are based GP model is effective for the electricity load data set.

Since the number of components in $\mathcal{R}_1$ and $\mathcal{R}_2$ are unknown, we set $G = 3, 6, 9, 12, 15$ for the mix-GP and mix-GPFR models trained using the EM algorithm. For proposed algorithm and the SEM algorithm, G is set to be 15. Besides, P is set to be 30. The experimental results are described in Table 3. For $\mathcal{R}_1$ and $\mathcal{R}_2$, the RMSE of proposed algorithm is smaller than that of the SEM algorithm since the number of components given by the SEM algorithm is smaller than the optimal one. The clustering results are denoted in Fig. 4. On the left and right side of Fig. 4 are the clustering results of proposed algorithm on the training and test data sets, respectively. For $\mathcal{R}_1$ and $\mathcal{R}_2$, the numbers of components given by proposed algorithm are 13 and 11, respectively. As can be seen from Fig. 4, curves belonging to different components are basically different to a certain extent, that is, the clustering given by the algorithm are reasonable.

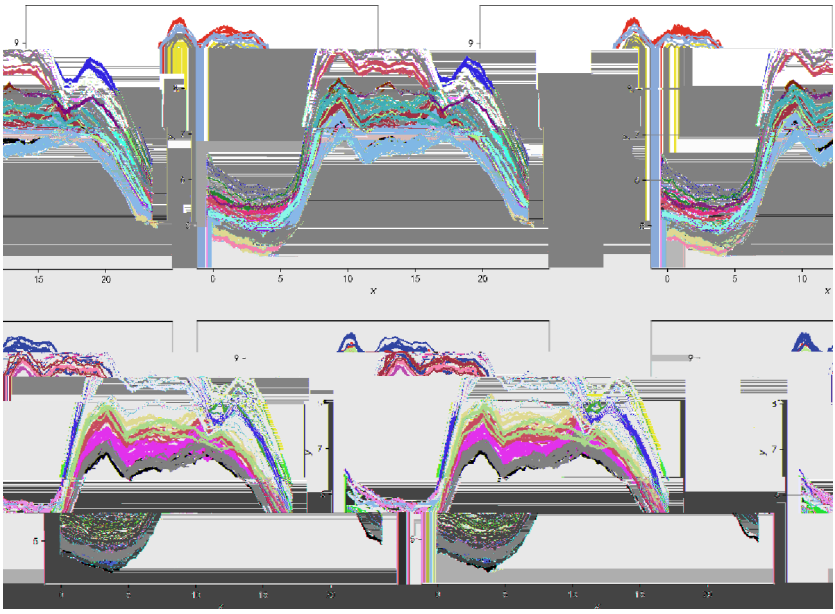

Fig. 4. Clustering results of fishery data set obtained automatic model selection algorithm $\mathcal{R}_1$ and $\mathcal{R}_2$.

Table 3. RMSE and training time of all the methods $\mathcal{R}_1$ and $\mathcal{R}_2$.

	$\mathcal{R}_1$		$\mathcal{R}_2$	
	RMSE	Time (min)	RMSE	Time (min)
GP	0.9599	19.43	0.8977	20.39
mix.-GP (3)	0.9390	20.33	0.8846	21.42
mix.-GP (6)	0.9387	22.54	0.8854	22.66
mix.-GP (9)	0.9380	25.99	0.8853	26.09
mix.-GP (12)	0.9395	29.83	0.8847	31.23
mix.-GP (15)	0.9401	34.76	0.8872	36.91
GPFR	0.5584	21.30	0.5499	21.59
mix.-GPFR (3)	0.2089	24.45	0.2133	20.64
mix.-GPFR (6)	0.1701	25.76	0.1731	24.77
mix.-GPFR (9)	0.1356	28.93	0.1455	29.45
mix.-GPFR (12)	0.1248	34.65	0.1314	33.63
mix.-GPFR (15)	0.1178	35.88	0.1301	36.78
mix.-GPFR (SEM)	0.1323	150.76	0.1377	170.17
mix.-GPFR (BYY)	**0.1109**	33.97	**0.1201**	34.58

5 Conclusion

I thi a e, e e a a t matic m del electi alg ithm ba ed BYY ha m lea i g f mix.-GPFR m del. Si ce diffe e t t ai i g c e ha e diffe e t i t, BYY ha m lea i g ca t be di ectl a lied t the m del electi blem f mix.-GPFR m del. T tackle thi , e e ec e ec t. cti ba ed GP m del, th gh hich, e if the i t f all the t ai i g c e. The , e ca make m del electi f mix.-GPFR m del ia BYY ha m lea i g. Ex. e i me tal e lt thetic a d eal- ld data et h that e ed a t matic m del electi alg ithm ca d the timal mbe f c m e t i a m lti- ce c e data et a d it time c m le.it i l e tha that f the SEM alg ithm.

Ackno ledgement. Thi k i ted b the Nati al Ke R & D P g am f Chi a (2018AAA0100205).

References

1. Ra m e , C.E., Williams, C.K.I.: Ga ia P ce e f Machi e Lea i g. MIT P e , Cambridge (2006)
2. Shi, J.Q., Wa g, B.: C e edicti a d cl te i g ith mix.t e f Ga ia ce f cti al eg e i m del. Stati t. C m t. **18**, 267 283 (2008)
3. Shi, J.Q., Ch i, T.: Ga ia P ce Reg e i A al i f F cti al Data. CRC P e , B ca Rat (2011)
4. W , D., Ma, J.: A DAEM alg ithm f mix.t e f Ga ia ce f cti al eg e i . I : H a g, D.-S., Ha , K., H ai , A. (ed .) ICIC 2016. LNCS (LNAI), l. 9773, . 294 303. S i ge , Cham (2016). htt ://d i. g/10.1007/978-3-319-42297-8_28
5. Qia g, Z., Ma, J.: A t matic m del electi f the mix.t e f Ga ia ce e f eg e i . I : H , X., Xia, Y., Zha g, Y., Zha , D. (ed .) ISNN 2015. LNCS, l. 9377, . 335 344. S i ge , Cham (2015). htt ://d i. g/10.1007/978-3-319-25393-0_37
6. Qia g, Z., L , J., Ma, J.: C e cl te i g ia the lit lea i g f mix.t e f Ga - ia ce e. I : 2016 IEEE 13th I te ati al C fe e ce Sig al P ce i g (ICSP), . 1089 1094 (2016)
7. Qia g, Z., Ma, J.: M del electi edicti f the mix.t e f Ga ia ce e ith RJMCMC. I : Shi, Z., Pe a t , C., H a g, T. (ed .) ICIS 2018. IAICT, l. 539, . 310 317. S i ge , Cham (2018). htt ://d i. g/10.1007/978-3-030-01313-4_33
8. Li, T., Ma, J.: Di ichlet ce mix.t e f Ga ia ce f cti al eg e i a d it a iati al EM alg ithm. Patte Rec g . **134**, 109 129 (2023)
9. Sch a , G.: E timati g the dime i f a m del. A . Stati t. **6**(2), 461 464 (1978)
10. Richa d , S., G ee , P.J.: O Ba e ia a al i f mix.t e ith a k mbe f c m e t. J R al Stati t. S c. (Se. B) **59**(4), 731 792 (1997)
11. E c bac, M.D., We t, M.: Ba e ia de it e timati a d i fe e ce i g mix.t e . J. Am. Stati t. A c. **90**(430), 577 588 (1995)
12. X , L.: Yi g-Ya g machi e: a Ba e ia -K llback cheme f i ed lea i g a d e ce lt ect a ti ati . I : P ceedi g f the 1995 I te ati al C fe e ce Ne al I f mati P ce i g, l. 2, . 977 988 (1995)
13. X , L.: Be t ha m , i ed RPCL a d a t mated m del electi f i ed a d i ed lea i g Ga ia mix.t e , thee-la e et a d ME-RBF-SVM m del. I t. J. Ne al S t. **11**(1), 43 69 (2001)

14. Che , G., Li, L., Ma, J.: A g adie t BYY ha rm lea i g alg ithm f t aight li e detecti . I : S , F., Zha g, J., Ta , Y., Ca , J., Y , W. (ed .) ISNN 2008. LNCS, l. 5263, . 618 626. S i ge , Heidelbe g (2008). htt ://d i. g/10.1007/978-3-540-87732-5_69
15. Li, L., Ma, J.: A BYY cale-i e me tal EM alg ithm f Ga ia mix.t e lea i g. A l. Math. C m t. **205**(2), 832 840 (2008)
16. Li, L., Ma, J.: A BYY lit-a d-me ge EM alg ithm f Ga ia mix.t e lea i g. I : S , F., Zha g, J., Ta , Y., Ca , J., Y , W. (ed .) ISNN 2008. LNCS, l. 5263, . 600 609. S i ge , Heidelbe g (2008). htt ://d i. g/10.1007/978-3-540-87732-5_67
17. Ma, J., Wa g, T., X , L.: A g adie t BYY ha rm lea i g le Ga ia mix.t e with a t rmated m del electi . Ne c m ti g **56**, 481 487 (2004)
18. Ma, J., Ga , B., Wa g, Y., Che g, Q.: C j gate a d at al g adie t le f BYY ha rm lea i g Ga ia mix.t e with a t rmated m del electi . I t. J. Patte Rec g . Arti f. I tell. **19**(5), 701 713 (2005)
19. Ma, J., Li , J.: The BYY a eali g lea i g alg ithm f Ga ia mix.t e with a t rmated m del electi . Patte Rec g **40**(7), 2029 2037 (2007)
20. Ma, J., He, X.: A fa t x.ed- i t BYY ha rm lea i g alg ithm Ga ia mix.t e with a t rmated m del electi . Patte Rec g . Lett. **29**(6), 701 711 (2008)
21. B , C.D.: O calc lati g with B- li e . J. A x.imati The . **6**, 50 62 (1972)