ABC-based Forecasting in State Space Models

Chaya Weerasinghe, Ruben Loaiza-Maya, Gael M. Martin and David T. Frazier

Department of Econometrics and Business Statistics
Monash University, Australia

Contact Email: chaya.w.patabedimuhamdiramalage@monash.edu

1. Exact Bayesian Forecasting in SSMs

- Observed data: $y_{1:T} = \{y_1, y_2 ..., y_T\}$
- Unobserved states: $x_{1:T} = \{x_1, x_2, ..., x_T\}$
- A SSM comprised of:
  - Measurement density: $p(y_t|x_t, \theta)$
  - (Markov) Transition density: $p(x_t|x_{t-1}, \theta)$
  - Initial state density: $p(x_1|\theta)$
- **Distribution of interest** is:

$$p(y_{T+1}|y_{1:T}) = \int_{\mathsf{X}}\int_{\mathsf{X}}\int_{\Theta} p(y_{T+1}|x_{T+1}, \theta) p(x_{T+1}|x_T, \theta) \times p(x_{1:T}|y_{1:T}, \theta) p(\theta|y_{1:T}) \, d\theta dx_{1:T} dx_{T+1}$$

- Often accessible via Bayesian MCMC methods or particle MCMC variants.

2. Intractable SSMs

- Occur in **two** ways:
  - When the **dimension** of either the $y_{1:T}$ or the $x_{1:T}$ (and the associated θ), or both, is **very large**
  - When the DGP is **unavailable** in closed form
- Can be analyzed using ***approximate computational methods:***
  - **Variational Bayes** (VB)
  - **Approximate Bayesian computation** (ABC) $\Leftarrow$ Our focus

3. Aim of ABC

- To produce draws from an **approximation** to $p(\theta|y_{1:T})$
- In the case where the DGP cannot be evaluated
- But *can* be **simulated** from

4. ABC accept/reject Algorithm (adapted for SSM)

- Simulate θ^i, $i = 1, 2, ..., N$, from $p(\theta)$.
- Simulate
  pseudo-states $x^s_{1:T}(\theta^i)$, $i = 1, 2, ..., N$ from $p(x^s_1|\theta^i)$ and $p(x^s_t|x^s_{t-1}, \theta^i)$, for $t = 2, ..., T$.
  pseudo-data $y^s_{1:T}(\theta^i)$, $i = 1, 2, ..., N$ from $p(y^s_t|x^s_t, \theta^i)$, for $t = 1, 2, ..., T$.
- Select $\theta^{(i)}$ such that:

$$d\{\eta(y_{1:T}), \eta\left(y^s_{1:T}(\theta^i)\right)\} \leq \varepsilon,$$

  $\eta(.)$ is a (vector) **summary statistic**
  $d\{.\}$ is a **distance criterion**
  the **tolerance** ε is arbitrarily small
- Selected draws $\Rightarrow$ draws from $p_\varepsilon(\theta|\eta(y_{1:T}))$

5. Objectives of this Research

- Explore the ***use of ABC*** to produce the probabilistic forecasts in SSMs
- Under ***misspecification*** of the DGP
- With the aim of producing accurate, fit for purpose predictions
- **Do not assume** correct model specification
- $\Rightarrow$ The whole focus of the Bayesian forecasting exercise **needs to change**
- Argument put forward in recent forecasting work:
  - **'Loss-based' or 'focused' Bayesian prediction**
  - **[Loaiza-Maya et al. (2021), Frazier et al. (2021)]**
- ***For the first time:*** ABC-based forecasting + loss-based prediction

6. Loss-based Bayesian Prediction

- In the spirit of the various generalized Bayesian *inferential* methods, **[Bissiri, Holmes and Walker(2016), Loaiza-Maya et al. (2021), Frazier et al. (2021)]**
- Replace the **likelihood function (log-score loss)** in the conventional Bayesian update
- By the *particular* **predictive loss** that matters for the *particular* forecasting problem being tackled

7. Loss-based Bayesian Prediction in SSMs

- $\mathcal{P}^{(T)} \Rightarrow$ Class of plausible predictive SSMs for Y_{T+1}
- Assumption: Model is ***misspecified***
- Construct the **loss-based posterior/Gibbs posterior** $p_L(\theta|y_{1:T})$:
  via some (positively-oriented) **scoring rule**

$$S_T(\theta) := \sum_{t=0}^{T-1} s(P^{(t)}_\theta, y_{t+1})$$

- **Loss-based predictive:**

$$p_L(y_{T+1}|y_{1:T}) = \int_{\mathsf{X}}\int_{\mathsf{X}}\int_{\Theta} p(y_{T+1}|x_{T+1}, \theta, y_{1:T}) p(x_{T+1}|x_T, \theta) \times p(x_{1:T}|y_{1:T}, \theta) p_L(\theta|y_{1:T}) d\theta dx_{1:T} dx_{T+1}$$

- **Immediate Problems???**

$$S_T(\theta) := \textstyle\sum_{t=0}^{T-1} s(\underbrace{p(y_{t+1}|y_{1:T}, \theta)}_{\text{marginal of } x_{1:T}}, y_{t+1})$$

Intractable SSM model $\Rightarrow$ $p_L(\theta|y_{1:T})$ out of reach

8. Loss-based Approximate Bayesian Forecasting (ABF)

- In the spirit of ***'auxiliary model'***-based ABC **[Drovandi et al., 2011]**,:
- We implement ABC by producing $\eta(\cdot)$, by maximizing

$$S_T(\beta) := \sum_{t=0}^{T-1} s(P^{(t)}_\beta, y_{t+1}) \quad \text{w.r.t. } \beta$$

- ***Summary statistics*** :
  Average of the first-derivative of $S_T(\beta)$ based on a scoring rule s_j evaluated at $\hat{\beta}_j(y_{1:T})$:

$$\eta_j(y^s_{1:T}) = \bar{S}_j\left\{y_{1:T}; \hat{\beta}_j(y_{1:T})\right\} = T^{-1} \frac{\partial \sum_{t=0}^{T-1} s_j(P^{(t)}_\beta, y^s_{t+1}(\theta^i))}{\partial \beta}\Bigg|_{\beta=\hat{\beta}_j(y_{1:T})}$$

- ***Distance Criterion***: Mahalanobis distance

$$d\{\eta\left(y^s_{1:T}\right), \eta\left(y_{1:T}\right)\} = \sqrt{\left[\bar{S}_j\left\{y^s_{1:T}(\theta^i); \hat{\beta}_j(y_{1:T})\right\}\right]' \hat{\Sigma} \left[\bar{S}_j\left\{y^s_{1:T}(\theta^i); \hat{\beta}_j(y_{1:T})\right\}\right]},$$

- **Loss-based ABC predictive:**

$$g_L(y_{T+1}|y_{1:T}) = \int_{\mathsf{X}}\int_{\mathsf{X}}\int_{\Theta} p(y_{T+1}|x_{T+1}, \theta) p(x_{T+1}|x_T, \theta) p(x_T|\theta, y_{1:T}) p_{L,\varepsilon}(\theta|\eta(y_{1:T})) d\theta dx_T dx_{T+1}.$$

  $\Rightarrow$ Gives higher predictive accuracy in the chosen s_j
- **A comparator:**
  - Using the auxiliary model predictive *directly* in a generalized Bayesian update
  - $\Rightarrow$ Focused Bayesian Prediction (FBP) **[Loaiza-Maya et al. (2021)]**

9. Numerical Illustration

- **Loss-based ABF**
  - Assumed DGP: **Stochastic volatility (SV) model** (an SSM)
  - True DGP1: same as assumed DGP $\Rightarrow$ Correct model specification
  - True DGP2: SV + asymmetry $\Rightarrow$ Model misspecification
  - Auxiliary models: ARCH(1), GARCH(1,1)
- **FBP**
  - Assumed DGP: same auxiliary model used for comparable loss-based ABF
  - True DGPs same as above
  - $\Rightarrow$ Always misspecified
- **Scoring rule** s_j
  - Log score (LS), Censored log score (CLS), Continuously ranked probability score (CRPS), Interval Score (IS)
- **Exact predictive**: accessed via Hybrid Gibbs-MH MCMC algorithm

10. Empirical Illustration

- Forecast daily returns on the S&P500 index
- Assumed DGP: SV model with $\alpha-$ stable transition $\Rightarrow$ ***truly intractable***

11. Results (based on GARCH(1,1) auxiliary model) and Conclusions

Table 1: **Predictive accuracy of loss-based ABF & FBP: Correct model specification of the SSM**

Panel A: Loss-based ABF — Average out-of-sample score

Scoring rule	LS	CLS10	CLS20	CLS80	CLS90	CRPS	IS
ABC-LS	*-0.8189*	*-0.2956*	-0.4842	*-0.4875*	*-0.2985*	**-0.3211**	-3.0575
ABC-CLS10	-0.8244	-0.2964	-0.4857	-0.4907	-0.3016	*-0.3216*	-3.0946
ABC-CLS20	-0.8193	*-0.2956*	*-0.4841*	-0.4880	-0.2990	**-0.3211**	-3.0595
ABC-CLS80	-0.8192	-0.2957	**-0.4840**	-0.4879	-0.2988	**-0.3211**	-3.0595
ABC-CLS90	-0.8237	-0.2986	-0.4875	-0.4885	-0.2993	*-0.3216*	-3.0932
ABC-CRPS	**-0.8188**	**-0.2955**	*-0.4841*	**-0.4874**	**-0.2983**	**-0.3211**	**-3.0551**
ABC-IS	**-0.8188**	**-0.2955**	**-0.4840**	-0.4876	-0.2986	**-0.3211**	*-3.0559*
Exact	-0.8191	-0.2957	-0.4842	-0.4876	-0.2986	-0.3211	-3.0594

Panel B: FBP — Average out-of-sample score

Scoring rule	LS	CLS10	CLS20	CLS80	CLS90	CRPS	IS
FBP-LS	*-0.8557*	-0.3119	-0.5020	-0.5009	-0.3111	*-0.3238*	*-3.1525*
FBP-CLS10	-0.8989	**-0.3046**	**-0.4934**	-0.5425	-0.3380	-0.3363	-3.2309
FBP-CLS20	-0.8838	*-0.3054*	*-0.4941*	-0.5287	-0.3274	-0.3322	-3.1963
FBP-CLS80	-0.8862	-0.3273	-0.5315	*-0.4959*	*-0.3061*	-0.3328	-3.2051
FBP-CLS90	-0.9154	-0.3470	-0.5579	**-0.4958**	**-0.3051**	-0.3408	-3.2699
FBP-CRPS	-0.8616	-0.3165	-0.5057	-0.5055	-0.3162	**-0.3228**	-3.1930
FBP-IS	**-0.8551**	-0.3099	-0.5011	-0.4997	-0.3094	-0.3242	**-3.1434**
Exact	-0.8191	-0.2957	-0.4842	-0.4876	-0.2986	-0.3211	-3.0594

Table 2: **Predictive accuracy of loss-based ABF & FBP: Misspecification of the SSM**

Panel A: Loss-based ABF — Average out-of-sample score

Scoring rule	LS	CLS10	CLS20	CLS80	CLS90	CRPS	IS
ABC-LS	*-1.3427*	-0.3586	**-0.6173**	-0.4900	-0.2975	*-0.5331*	-4.6333
ABC-CLS10	-1.4117	*-0.3572*	-0.6327	-0.5122	-0.3037	-0.5616	*-4.5813*
ABC-CLS20	-1.3737	**-0.3553**	*-0.6202*	-0.5062	-0.3084	-0.5427	-4.7791
ABC-CLS80	-2.0917	-0.8118	-1.2925	**-0.4675**	*-0.2822*	-0.6082	-10.5000
ABC-CLS90	-2.4259	-0.8961	-1.4896	*-0.4715*	**-0.2777**	-0.6509	-12.7820
ABC-CRPS	**-1.3371**	-0.3629	-0.6214	-0.4881	-0.2998	**-0.5309**	-4.7405
ABC-IS	-1.4882	-0.3657	-0.6648	-0.5333	-0.3057	-0.6025	**-4.2895**
Exact	-1.3343	-0.3618	-0.6199	-0.4882	-0.3003	-0.5304	-4.7357

Panel B: FBP — Average out-of-sample score

Scoring rule	LS	CLS10	CLS20	CLS80	CLS90	CRPS	IS
FBP-LS	**-1.3471**	-0.3694	-0.6312	-0.4840	-0.2954	*-0.5340*	*-4.6822*
FBP-CLS10	-1.3669	**-0.3593**	*-0.6243*	-0.5094	-0.3212	-0.5380	-4.8663
FBP-CLS20	-1.4076	*-0.3601*	**-0.6201**	-0.5606	-0.3749	-0.5464	-5.2801
FBP-CLS80	-2.0718	-0.9227	-1.3493	**-0.4491**	*-0.2657*	-0.5888	-8.9718
FBP-CLS90	-2.5938	-1.1223	-1.7430	*-0.4579*	**-0.2644**	-0.6357	-11.5340
FBP-CRPS	*-1.3485*	-0.3786	-0.6378	-0.4839	-0.2973	**-0.5319**	-4.7954
FBP-IS	-1.4521	-0.3734	-0.6765	-0.5106	-0.2943	-0.5979	**-4.3179**
Exact	-1.3343	-0.3618	-0.6199	-0.4882	-0.3003	-0.5304	-4.7357

Table 3: **Predictive accuracy of loss-based ABF & FBP: Empirical Illustration**

Panel A: Loss-based ABF — Average out-of-sample score

Scoring rule	LS	CLS10	CLS20	CLS80	CLS90
ABC-LS	3.3967	0.0268	0.2923	0.5170	0.1523
ABC-CLS10	**3.4058**	**0.0391**	**0.3048**	0.5122	0.1500
ABC-CLS20	*3.4053*	*0.0383*	*0.3041*	0.5126	0.1502
ABC-CLS80	3.3772	0.0052	0.2713	**0.5192**	**0.1547**
ABC-CLS90	3.3849	0.0134	0.2790	*0.5191*	*0.1540*

Panel B: FBP — Average out-of-sample score

Scoring rule	LS	CLS10	CLS20	CLS80	CLS90
FBP-LS	**3.3467**	-0.0262	0.2406	0.4801	0.1473
FBP-CLS10	3.1051	**0.0463**	**0.3096**	0.1903	-0.1119
FBP-CLS20	3.1788	*0.0417*	*0.3085*	0.2645	-0.0430
FBP-CLS80	*3.3293*	-0.0515	0.2155	**0.4860**	*0.1531*
FBP-CLS90	3.1941	-0.1204	0.1042	*0.4820*	**0.1543**

- Positively-oriented scores $\Rightarrow$ large (**in bold**) is good

Correct model specification (Table 1):
- Panel A: focusing reaps no benefit
- Loss-based ABF $\equiv$ **misspecified exact Bayes**
- Panel B: ***'coherent'*** predictions are in evidence
- **Loss-based ABF > FBP**

Model misspecification (Table 2):
- Looking for **bold** values on the diagonal $\Rightarrow$ ***'coherent'*** predictions in evidence
- **Loss-based ABF > misspecified exact Bayes**
- **Comparison**: Dominance of loss-based ABF over FBP is **not uniform**

Empirical illustration (Table 3):
- Conclusions are a bit mixed
- But overall, **loss-based ABF is more accurate than FBP**
- Often markedly so