ABC-based Forecasting in State Space Models

Chaya Weerasinghe
Department of Econometrics and Business Statistics
Monash University

IMS-APRM 2024

Supervisors:
Professor Gael M. Martin, Dr. Ruben Loaiza Maya and Associate Professor David Frazier

State Space Models

- Powerful tool for modelling and forecasting the complex dynamics associated with the observed time series data

State Space Models

- Powerful tool for modelling and forecasting the complex dynamics associated with the observed time series data
- Observed data: $y_{1:T} = \{y_1, y_2 ..., y_T\}$
- Unobserved states: $x_{1:T} = \{x_1, x_2, ..., x_T\}$

State Space Models

- Powerful tool for modelling and forecasting the complex dynamics associated with the observed time series data
- Observed data: $y_{1:T} = \{y_1, y_2 ..., y_T\}$
- Unobserved states: $x_{1:T} = \{x_1, x_2, ..., x_T\}$
- A state space model (SSM) comprised of two parts:

State Space Models

- Powerful tool for modelling and forecasting the complex dynamics associated with the observed time series data
- Observed data: $y_{1:T} = \{y_1, y_2 ..., y_T\}$
- Unobserved states: $x_{1:T} = \{x_1, x_2, ..., x_T\}$
- A state space model (SSM) comprised of two parts:

Measurement density: $p(y_t|x_t, \theta)$

State Space Models

- Powerful tool for modelling and forecasting the complex dynamics associated with the observed time series data
- Observed data: $y_{1:T} = \{y_1, y_2 ..., y_T\}$
- Unobserved states: $x_{1:T} = \{x_1, x_2, ..., x_T\}$
- A state space model (SSM) comprised of two parts:

$$\begin{aligned} &\text{Measurement density:} && p(y_t | x_t, \theta) \\ &\text{(Markov) transition density:} && p(x_t | x_{t-1}, \theta) \end{aligned}$$

State Space Models

- Powerful tool for modelling and forecasting the complex dynamics associated with the observed time series data
- Observed data: $y_{1:T} = \{y_1, y_2..., y_T\}$
- Unobserved states: $x_{1:T} = \{x_1, x_2, ..., x_T\}$
- A state space model (SSM) comprised of two parts:

$$\text{Measurement density:} \qquad p(y_t | x_t, \theta)$$

$$\text{(Markov) transition density:} \quad p(x_t | x_{t-1}, \theta)$$

- Plus

$$\text{Initial state density:} \quad p(x_1 | \theta)$$

- May depend on a set of unknown, **static** parameters θ

Exact Bayesian Forecasting in SSMs

- **Distribution of interest** is

$$p(y_{T+1}|y_{1:T})$$

Exact Bayesian Forecasting in SSMs

- **Distribution of interest** is

$$
\begin{aligned}
p(y_{T+1}|y_{1:T}) \quad = \quad & \int_{\mathsf{X}}\int_{\mathsf{X}}\int_{\Theta} p(y_{T+1}|x_{T+1},\theta)p(x_{T+1}|x_T,\theta) \\
& \times \underbrace{\left(p(x_{1:T}|y_{1:T},\theta)\right)} \underbrace{p\left(\theta|y_{1:T}\right)} d\theta dx_{1:T} dx_{T+1}
\end{aligned}
$$

Exact Bayesian Forecasting in SSMs

- **Distribution of interest** is

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

- Often readily accessible via
  - Bayesian Markov chain Monte Carlo (MCMC) methods or
  - Particle MCMC variants
- But... challenges remain when the model is **intractable** in some sense

Intractable SSMs

- Occurs in **two** ways:

Intractable SSMs

- Occurs in **two** ways:

  1. when the **dimension** of either **y** or **x** (and the associated θ), or both, is **very large**

Intractable SSMs

- Occurs in **two** ways:
  1. when the **dimension** of either $\mathbf{y}$ or $\mathbf{x}$ (and the associated θ), or both, is **very large**
  2. when the data generating process (DGP) **unavailable** in closed form

Intractable SSMs

- Occurs in **two** ways:
    1. when the **dimension** of either $\mathbf{y}$ or $\mathbf{x}$ (and the associated θ), or both, is **very large**
    2. when the data generating process (DGP) **unavailable** in closed form
- $\Rightarrow$ exact Bayesian prediction **may not be feasible**

Intractable SSMs

- Can be analyzed using ***approximate computational methods***

Intractable SSMs

- Can be analyzed using ***approximate computational methods***
- High-dimensional SSMs are tackled via **variational Bayes** (VB) methods :
  - **[Tran et al., 2017], [Quiroz et al., 2022], [Loaiza-Maya et al., 2021b], [Frazier et al., 2023]**

Intractable SSMs

- Can be analyzed using ***approximate computational methods***

- High-dimensional SSMs are tackled via **variational Bayes** (VB) methods :
  - **[Tran et al., 2017], [Quiroz et al., 2022], [Loaiza-Maya et al., 2021b], [Frazier et al., 2023]**

- SSMs with unavailable components are managed via **approximate Bayesian computation** (ABC) :
  - **[Dean et al., 2014], [Creel and Kristensen, 2015], [Frazier et al., 2019], [Martin et al., 2019]**
  - $\Rightarrow$ **Our focus**

We focus on ABC...

- The **aim of ABC**
  - is to produce draws from an **approximation** to $p(\theta|y_{1:T})$
  - in the case where DGP cannot be evaluated
  - But *can* be **simulated** from

ABC Algorithm (adapted for SSM)

ABC Algorithm (adapted for SSM)

1. Simulate $i = 1, 2, ..., N$, $i.i.d.$ draws of $\theta^{(i)}$ from $p(\theta)$

ABC Algorithm (adapted for SSM)

1. Simulate $i = 1, 2, ..., N$, $i.i.d.$ draws of $\theta^{(i)}$ from $p(\theta)$
2. 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 $t = 1, 2, ..., T$.

ABC Algorithm (adapted for SSM)

1. Simulate $i = 1, 2, ..., N$, $i.i.d.$ draws of $\theta^{(i)}$ from $p(\theta)$
2. 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 $t = 1, 2, ..., T$.
3. Select $\theta^{(i)}$ such that:

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

ABC Algorithm (adapted for SSM)

1. Simulate $i = 1, 2, ..., N$, *i.i.d.* draws of $\theta^{(i)}$ from $p(\theta)$
2. 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 $t = 1, 2, ..., T$.
3. 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

ABC Algorithm (adapted for SSM)

1. Simulate $i = 1, 2, ..., N$, $i.i.d.$ draws of $\theta^{(i)}$ from $p(\theta)$
2. 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 $t = 1, 2, ..., T$.
3. 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
4. Selected draws $\Rightarrow$ draws from $p_\varepsilon(\theta|\eta(y_{1:T}))$

Existing ABC Literature

Existing ABC Literature

- Under **correct model specification,**
  - ABC posterior - Bayesian consistent for the true parameter, asymptotically normal
    [Frazier et al., 2018] - 'Asymptotic properties of approximate Bayesian computation'
  - also in an explicitly SSM setting
    [Martin et al., 2019] - 'Auxiliary likelihood-based approximate Bayesian computation in state space models'
  - ABC-based forecasting
    [Frazier et al., 2019] - 'Approximate Bayesian forecasting'

Existing ABC Literature

- Under **correct model specification,**
  - ABC posterior - Bayesian consistent for the true parameter, asymptotically normal
    [Frazier et al., 2018] - 'Asymptotic properties of approximate Bayesian computation'
  - also in an explicitly SSM setting
    [Martin et al., 2019] - 'Auxiliary likelihood-based approximate Bayesian computation in state space models'
  - ABC-based forecasting
    [Frazier et al., 2019] - 'Approximate Bayesian forecasting'
- Under **model misspecification**
  - **[Frazier et al., 2020] - 'Model misspecification in approximate Bayesian computation: consequences and diagnostics'**

What is left unexplored?

- **ABC-based forecasting** – with *misspecified* SSMs

What is left unexplored?

- **ABC-based forecasting** – with *misspecified* SSMs
- If we *do not assume* the correct model specification $\Rightarrow$
- the whole focus of the Bayesian forecasting exercise **needs to change**.
  - argument put forward in recent forecasting works by
    - **[Loaiza-Maya et al., 2021a] - 'Focused Bayesian prediction'**
    - **[Frazier et al., 2021] - 'Loss-based variational Bayes prediction'**

What is left unexplored?

- **ABC-based forecasting** – with ***misspecified*** SSMs
- If we ***do not assume*** the correct model specification ⇒
- the whole focus of the Bayesian forecasting exercise **needs to change**.
    - argument put forward in recent forecasting works by
        - **[Loaiza-Maya et al., 2021a] - 'Focused Bayesian prediction'**
        - **[Frazier et al., 2021] - 'Loss-based variational Bayes prediction'**
- *For the first time:* **ABC-based forecasting + Loss-based Bayesian prediction**

Loss-based Bayesian Prediction

- **Essence of the idea:**

Loss-based Bayesian Prediction

- **Essence of the idea:**
- In the spirit of the various generalized Bayesian *inferential* methods,
    - [Bissiri et al., 2016]; [Giummolè et al., 2019]; [Knoblauch et al., 2019]; [Pacchiardi and Dutta, 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

Loss-based Bayesian Prediction in SSMs

- Assume a class of plausible **predictive SSMs** for Y_{T+1}, conditioned on information $\mathcal{F}_T$:

$$\mathcal{P}^{(T)} := \{P_\theta^{(T)}, \theta \in \Theta\}$$

- **Key thing:** Do not assume $P_0 \in \mathcal{P}^{(T)}$

- Construct the **loss-based posterior/Gibbs posterior** $p_L(\theta|y_{1:T})$
  - via some (positively-oriented) **scoring rule**

$$S_T(\theta) := \sum_{t=1}^{T} s(P_\theta^{(t)}, y_{t+1})$$

Loss-based Bayesian Prediction in SSMs

- $\Rightarrow$ **Loss-based predictive**

$$
\begin{aligned}
& p_L(y_{T+1}|y_{1:T}) \\
= \; & \int_{\mathsf{X}}\int_{\mathsf{X}}\int_{\Theta} p(y_{T+1}|x_{T+1},\theta)p(x_{T+1}|x_T,\theta) \\
& \times p(x_T|y_{1:T},\theta)\color{red}{p_L(\theta|y_{1:T})}\color{black}{d\theta dx_{1:T}dx_{T+1}}
\end{aligned}
$$

Immediate Problems...

- Consider the **loss-based predictive** $p_L(y_{T+1}|y_{1:T})$ which is constructed based on $p_L(\theta|y_{1:T})$ via

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

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

Our solution..

- Approximate $p_L(\theta|y_{1:T})$ using **ABC**

Our solution..

- Approximate $p_L(\theta|y_{1:T})$ using **ABC**
  - In the spirit of ***'auxiliary model'***- based ABC :
    - **[Drovandi et al., 2011], [Drovandi et al., 2015] and [Martin et al., 2019]**
  - We choose an **auxiliary** model
    - that is a 'reasonable' approximation to the assumed **SSM** &
    - admit a **closed-form predictive** (with parameter vector β)

Our solution..

- Approximate $p_L(\theta|y_{1:T})$ using **ABC**
  - In the spirit of ***'auxiliary model'***- based ABC :
    - **[Drovandi et al., 2011], [Drovandi et al., 2015] and [Martin et al., 2019]**
  - We choose an **auxiliary** model
    - that is a 'reasonable' approximation to the assumed **SSM** &
    - admit a **closed-form predictive** (with parameter vector β)
  - Define the score-based criterion as:

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

  - **NOTE:** this is useful ***even when the DGP itself is tractable***

How Do We Choose the Summary Statistics?

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

How Do We Choose the Summary Statistics?

- *Summary statistics* : average of the first-derivative of $S_T(\beta)$ using a scoring rule s_j
  - evaluated at $\hat{\beta}_j(y_{1:T})$: optimizer of $S_T(\beta)$ using s_j
- That is,

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

and:

$$\eta_j(y_{1:T}) = T^{-1} \left. \frac{\partial \sum_{t=1}^{T} s_j(P^{(t)}_{\beta}, y_{t+1})}{\partial \beta} \right|_{\beta=\hat{\beta}_j(y_{1:T})} = 0$$

Distance Criterion

- ***ABC distance*** : Mahalanobis distance

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

  - where $\hat{\Sigma}$ is the inverse of the (estimated) covariance matrix of $\eta_j(y^s_{1:T})$ across draws and
  - $\bar{S}_j\left\{y^s_{1:T}(\theta^i);\hat{\beta}_j(y_{1:T})\right\} = T^{-1}\frac{\partial\sum_{t=1}^{T}s_j(P_\beta^{(t)},y^s_{t+1}(\theta^i))}{\partial\beta}\Big|_{\beta=\hat{\beta}_j(y_{1:T})}$

Loss-based ABF

- The accepted draws of θ produced by the ABC algorithm $\Rightarrow$
- are the draws from **loss-based ABC posterior**, $p_{L,\varepsilon}(\theta|\eta(y_{1:T}))$

Loss-based ABF

- The accepted draws of θ produced by the ABC algorithm $\Rightarrow$
- are the draws from **loss-based ABC posterior**, $p_{L,\varepsilon}(\theta|\eta(y_{1:T}))$
- **Loss-based ABC predictive** :

$$
\begin{aligned}
& g_L(y_{T+1}|y_{1:T}) \\
= \; & \int_{\mathsf{X}}\int_{\mathsf{X}}\int_{\theta} p(y_{T+1}|x_{T+1},\theta)p(x_{T+1}|x_T,\theta)p(x_T|\theta,y_{1:T}) \\
\times \; & p_{L,\varepsilon}(\theta|\eta(y_{1:T}))d\theta dx_T dx_{T+1}
\end{aligned}
$$

- $\Rightarrow$ loss-based ABC forecasting (**loss-based ABF**)

Comparator to Loss-based ABF

- Using the **auxiliary model predictive** ***directly*** in a generalized Bayesian update
  - Avoids the use of a SSM representation
  - But, use of a *simpler model* in a focused up-date well be adequate
  - ⇒ **[Loaiza-Maya et al., 2021a] - 'Focused Bayesian prediction' (FBP)**

Numerical Illustration: Simulation design

- **Predictive class**, $P_\theta^{(t)}$ is a SSM:
  - **SV model** for a continuous financial return, y_t,

$$y_t = \mu + e^{\alpha_t/2} e_t \quad ; \quad e_t \sim N(0,1)$$

$$\alpha_t = \bar{h}_\alpha + \phi(\alpha_{t-1} - \bar{h}_\alpha) + w_t \quad ; \quad w_t \sim N(0, \sigma_\alpha^2)$$

$$\alpha_0 \sim N\left(\bar{h}_\alpha, \frac{\sigma_\alpha^2}{1-\phi^2}\right),$$

- $\theta = (\phi, \sigma_\alpha^2, \mu, \bar{h}_\alpha)'$

Numerical Illustration: Simulation design

- **True DGP**

Numerical Illustration: Simulation design

- **True DGP**
  1. A model that matches the assumed SV model
     - $\Rightarrow$ **Correct model specification**

Numerical Illustration: Simulation design

- **True DGP**
  1. A model that matches the assumed SV model
     - $\Rightarrow$ **Correct model specification**
  2. SV model that better replicates the stylized features of financial returns data, as used by **[Loaiza-Maya et al., 2021a]** :

$$z_t = e^{h_t/2}\epsilon_t$$
$$h_t = \bar{h} + a(h_{t-1} - \bar{h}) + \sigma_h \eta_t$$
$$y_t = G^{-1}(F_z(z_t))$$

     - $\Rightarrow$ Implied copula of a **stochastic volatility** model combined with a **skewed normal marginal,** $g(y_t)$ (imposed via G^{-1})
     - $\Rightarrow$ **Model misspecification**
- Predicting **extreme returns** accurately is important
- Will **focus** on that goal $\Rightarrow$ use an appropriate s_j in the update

Numerical Illustration: Simulation design

- Two **auxiliary models** are used:
  1. a Gaussian ARCH(1) model
  2. a Gaussian GARCH(1,1) model

Numerical Illustration: Simulation design

- Two **auxiliary models** are used:
  1. a Gaussian ARCH(1) model
  2. a Gaussian GARCH(1,1) model

- Results for three types of **scores** reproduced here:
  1. Log score (LS)
  2. **Censored** log score (CLS) (rewards **predictive accuracy in a tail**)
  3. Continuously ranked probability score (CRPS) (rewards the assignment of high predictive mass near to the realized value)

  - **[Gneiting and Raftery, 2007]**

Numerical Illustration: Simulation design

- Two **auxiliary models** are used:
  1. a Gaussian ARCH(1) model
  2. a Gaussian GARCH(1,1) model

- Results for three types of **scores** reproduced here:
  1. Log score (LS)
  2. **Censored** log score (CLS) (rewards **predictive accuracy in a tail**)
  3. Continuously ranked probability score (CRPS) (rewards the assignment of high predictive mass near to the realized value)
     - [Gneiting and Raftery, 2007]

- Will discuss the results based on **Gaussian GARCH(1,1)** model under **model misspecification**

Simulation Results: Model misspecification

Loss-based ABF

Average out-of-sample score

ABC Score	LS	$CLS_{<10\%}$	$CLS_{>80\%}$	$CLS_{>90\%}$	CRPS
LS	-1.3427	-0.3586	-0.4900	-0.2975	-0.5331
$CLS_{<10\%}$	-1.4117	-0.3572	-0.5122	-0.3037	-0.5616
$CLS_{>80\%}$	-2.0917	-0.8118	-0.4675	-0.2822	-0.6082
$CLS_{>90\%}$	-2.4259	-0.8961	-0.4715	-0.2777	-0.6509
CRPS	-1.3371	-0.3629	-0.4881	-0.2998	-0.5309
Exact Bayes	-1.3343	0.3618	-0.4882	-0.3003	-0.5304

- **Rows** $\Rightarrow$ Scoring rules used in underlying loss-based ABF results
- **Last Row** $\Rightarrow$ Exact (but misspecified) Bayesian predictive
- **Columns** $\Rightarrow$ Measure of out-of-sample accuracy

Simulation Results: Model misspecification

Loss-based ABF

Average out-of-sample score

ABC Score	LS	$CLS_{<10\%}$	$CLS_{>80\%}$	$CLS_{>90\%}$	CRPS
LS	*-1.3427*	-0.3586	-0.4900	-0.2975	-0.5331
$CLS_{<10\%}$	-1.4117	*-0.3572*	-0.5122	-0.3037	-0.5616
$CLS_{>80\%}$	-2.0917	-0.8118	**-0.4675**	-0.2822	-0.6082
$CLS_{>90\%}$	-2.4259	-0.8961	-0.4715	**-0.2777**	-0.6509
CRPS	-1.3371	-0.3629	-0.4881	-0.2998	**-0.5309**
Exact Bayes	-1.3343	0.3618	-0.4882	-0.3003	-0.5304

- Positively-oriented scores ⇒ large (**in bold**) is good
- Looking for **bold** values on the diagonal
  - The predictive constructed via the use of a particular scoring rule predicts best according to that rule
  - **'Coherent predictions'** are in evidence

Simulation Results: Model misspecification

Loss-based ABF
Average out-of-sample score

ABC Score	LS	$CLS_{<10\%}$	$CLS_{>80\%}$	$CLS_{>90\%}$	CRPS
LS	*-1.3427*	-0.3586	-0.4900	-0.2975	-0.5331
$CLS_{<10\%}$	-1.4117	*-0.3572*	-0.5122	-0.3037	-0.5616
$CLS_{>80\%}$	-2.0917	-0.8118	**-0.4675**	-0.2822	-0.6082
$CLS_{>90\%}$	-2.4259	-0.8961	-0.4715	**-0.2777**	-0.6509
CRPS	-1.3371	-0.3629	-0.4881	-0.2998	**-0.5309**
Exact Bayes	-1.3343	**-0.3618**	**-0.4882**	**-0.3003**	**-0.5304**

- **Loss-based ABF $>$ misspecified exact Bayes**

Simulation Results: Model misspecification

Comparison of loss-based ABF and FBP

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

Scoring rule	LS	$\text{CLS}_{<10\%}$	$\text{CLS}_{>80\%}$	$\text{CLS}_{>90\%}$	CRPS
ABC-LS	*-1.3427*	-0.3586	-0.4900	-0.2975	*-0.5331*
ABC-$\text{CLS}_{<10\%}$	-1.4117	*-0.3572*	-0.5122	-0.3037	-0.5616
ABC-$\text{CLS}_{>80\%}$	-2.0917	-0.8118	**-0.4675**	*-0.2822*	-0.6082
ABC-$\text{CLS}_{>90\%}$	-2.4259	-0.8961	*-0.4715*	**-0.2777**	-0.6509
ABC-CRPS	**-1.3371**	-0.3629	-0.4881	-0.2998	**-0.5309**
Exact	-1.3343	-0.3618	-0.4882	-0.3003	-0.5304

Panel B: FBP
Average out-of-sample score

Scoring rule	LS	$\text{CLS}_{<10\%}$	$\text{CLS}_{>80\%}$	$\text{CLS}_{>90\%}$	CRPS
FBP-LS	**-1.3471**	-0.3694	-0.4840	-0.2954	*-0.5340*
FBP-$\text{CLS}_{<10\%}$	-1.3669	**-0.3593**	-0.5094	-0.3212	-0.5380
FBP-$\text{CLS}_{>80\%}$	-2.0718	-0.9227	**-0.4491**	*-0.2657*	-0.5888
FBP-$\text{CLS}_{>90\%}$	-2.5938	-1.1223	*-0.4579*	**-0.2644**	-0.6357
FBP-CRPS	*-1.3485*	-0.3786	-0.4839	-0.2973	**-0.5319**
Exact	-1.3343	-0.3618	-0.4882	-0.3003	-0.5304

Simulation Results: Model misspecification

Comparison of loss-based ABF and FBP

- Both methods produce ***coherent*** predictions
- Dominance of loss-based ABF over FBP is ***not uniform***
  - Likely to depend on the extent to which assumed SSM is misspecified

Empirical Illustration

- Applying **loss-based ABF** to daily returns data on the S & P500 index
- **Assumed predictive class**:
  - **SV model with α-Stable errors** for a continuous financial return, y_t,

$$y_t = e^{h_t/2}\varepsilon_t \quad ; \quad \varepsilon_t \overset{iid}{\sim} N(0,1)$$

$$h_t = \omega + \rho h_{t-1} + \sigma_h \eta_t \quad ; \quad \eta_t \overset{iid}{\sim} S(\alpha, -1, 0, dt = 1)$$

  - **Transition density is unavailable**
  - $\Rightarrow$ **Truly intractable SSM**
- $\theta = (\omega, \rho, \sigma_h^2, \alpha)'$

Empirical Illustration: Results

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

Scoring rule	LS	$CLS_{<10\%}$	$CLS_{<20\%}$	$CLS_{>80\%}$	$CLS_{>90\%}$
ABC-LS	3.3967	0.0268	0.2923	0.5170	0.1523
ABC-$CLS_{<10\%}$	**3.4058**	**0.0391**	**0.3048**	0.5122	0.1500
ABC-$CLS_{<20\%}$	*3.4053*	*0.0383*	*0.3041*	0.5126	0.1502
ABC-$CLS_{>80\%}$	3.3772	0.0052	0.2713	**0.5192**	**0.1547**
ABC-$CLS_{>90\%}$	3.3849	0.0134	0.2790	*0.5191*	*0.1540*

Panel B: FBP
Average out-of-sample score

Scoring rule	LS	$CLS_{<10\%}$	$CLS_{<20\%}$	$CLS_{>80\%}$	$CLS_{>90\%}$
FBP-LS	**3.3467**	-0.0262	0.2406	0.4801	0.1473
FBP-$CLS_{<10\%}$	3.1051	**0.0463**	**0.3096**	0.1903	-0.1119
FBP-$CLS_{<20\%}$	3.1788	*0.0417*	*0.3085*	0.2645	-0.0430
FBP-$CLS_{>80\%}$	*3.3293*	-0.0515	0.2155	**0.4860**	*0.1531*
FBP-$CLS_{>90\%}$	3.1941	-0.1204	0.1042	*0.4820*	**0.1543**

- Conclusions are a bit mixed
- But overall, loss-based ABF > FBP

Concluding remarks

- A new approach for conducting Bayesian prediction in **intractable** and ***misspecified*** SSMs
  - **Loss-based ABF**

Concluding remarks

- A new approach for conducting Bayesian prediction in **intractable** and ***misspecified*** SSMs
  - **Loss-based ABF**

- Two comparators are entertained:
  - Exact (but misspecified) Bayesian prediction
  - Prediction based on a generalized Bayesian up-date using auxiliary model alone - FBP

Concluding remarks

- A new approach for conducting Bayesian prediction in **intractable** and ***misspecified*** SSMs
  - **Loss-based ABF**
- Two comparators are entertained:
  - Exact (but misspecified) Bayesian prediction
  - Prediction based on a generalized Bayesian up-date using auxiliary model alone - FBP
- Simulation and Empirical results:
  - Coherent predictions are in evidence
  - More accurate forecasts than the exact (misspecified) predictive
  - Often more accurate than the FBP results

Concluding remarks

- A new approach for conducting Bayesian prediction in **intractable** and ***misspecified*** SSMs
  - **Loss-based ABF**
- Two comparators are entertained:
  - Exact (but misspecified) Bayesian prediction
  - Prediction based on a generalized Bayesian up-date using auxiliary model alone - FBP
- Simulation and Empirical results:
  - Coherent predictions are in evidence
  - More accurate forecasts than the exact (misspecified) predictive
  - Often more accurate than the FBP results
- Little to lose from adopting the ABC-based approach, and much to gain!

Thank you!

References I

Bissiri, P. G., Holmes, C. C., and Walker, S. G. (2016).
A general framework for updating belief distributions.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 78(5):1103–1130.

Creel, M. and Kristensen, D. (2015).
ABC of SV: Limited information likelihood inference in stochastic volatility jump-diffusion models.
Journal of Empirical Finance, 31:85–108.

Dean, T. A., Singh, S. S., Jasra, A., and Peters, G. W. (2014).
Parameter estimation for hidden Markov models with intractable likelihoods.
Scandinavian Journal of Statistics, 41(4):970–987.

Drovandi, C. C., Pettitt, A. N., and Faddy, M. J. (2011).
Approximate bayesian computation using indirect inference.
Journal of the Royal Statistical Society: Series C (Applied Statistics), 60(3):317–337.

Drovandi, C. C., Pettitt, A. N., and Lee, A. (2015).
Bayesian indirect inference using a parametric auxiliary model.
Statistical Science, 30(1):72–95.

Frazier, D. T., Loaiza-Maya, R., and Martin, G. M. (2023).
Variational bayes in state space models: Inferential and predictive accuracy.
Journal of Computational and Graphical Statistics, 32(3):793–804.

Frazier, D. T., Loaiza-Maya, R., Martin, G. M., and Koo, B. (2021).
Loss-based variational Bayes prediction.
arXiv preprint arXiv:2104.14054.

References II

Frazier, D. T., Maneesoonthorn, W., Martin, G. M., and McCabe, B. P. (2019).
Approximate bayesian forecasting.
International Journal of Forecasting, 35(2):521–539.

Frazier, D. T., Martin, G. M., Robert, C. P., and Rousseau, J. (2018).
Asymptotic properties of approximate bayesian computation.
Biometrika, 105(3):593–607.

Frazier, D. T., Robert, C. P., and Rousseau, J. (2020).
Model misspecification in approximate bayesian computation: consequences and diagnostics.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 82(2):421–444.

Giummolè, F., Mameli, V., Ruli, E., and Ventura, L. (2019).
Objective bayesian inference with proper scoring rules.
Test, 28(3):728–755.

Gneiting, T. and Raftery, A. E. (2007).
Strictly proper scoring rules, prediction, and estimation.
Journal of the American statistical Association, 102(477):359–378.

Knoblauch, J., Jewson, J., and Damoulas, T. (2019).
Generalized variational inference: Three arguments for deriving new posteriors.
arXiv preprint arXiv:1904.02063.

Loaiza-Maya, R., Martin, G. M., and Frazier, D. T. (2021a).
Focused bayesian prediction.
Journal of Applied Econometrics, 36(5):517–543.

References III

Loaiza-Maya, R., Smith, M. S., Nott, D. J., and Danaher, P. J. (2021b).
Fast and accurate variational inference for models with many latent variables.
Forthcoming. Journal of Econometrics.

Martin, G. M., McCabe, B. P., Frazier, D. T., Maneesoonthorn, W., and Robert, C. P. (2019).
Auxiliary likelihood-based approximate bayesian computation in state space models.
Journal of Computational and Graphical Statistics, 28(3):508–522.

Pacchiardi, L. and Dutta, R. (2021).
Generalized bayesian likelihood-free inference using scoring rules estimators.
arXiv preprint arXiv:2104.03889.

Quiroz, M., Nott, D. J., and Kohn, R. (2022).
Gaussian variational approximation for high-dimensional state space models.
https://arXiv:1801.07873.
Forthcoming, Bayesian Analysis.

Tran, M.-N., Nott, D. J., and Kohn, R. (2017).
Variational Bayes with intractable likelihood.
Journal of Computational and Graphical Statistics, 26(4):873–882.