A circular hidden Markov model for directional time series

Piumi Perera[1], Francis K.C. Hui[1] Carolyn Huston[2] A.H. Welsh[1]

[1]Research School of Finance, Actuarial Studies and Statistics
The Australian National University, Canberra, Australia

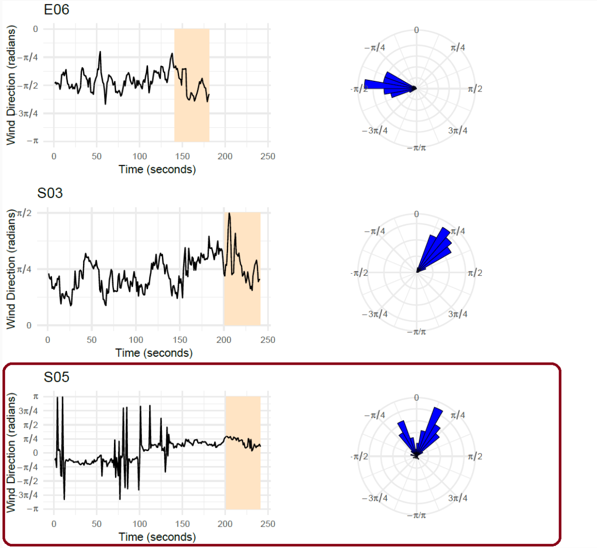
[2]CSIRO, Clayton South, Australia

Reference

Perera, A.A.P.N.M., Hui, F.K.C, Huston, C. and Welsh, A.H., (2025). *A circular hidden Markov model for directional time series data*. Japanese Journal of Statistics and Data Science.
https://doi.org/10.1007/s42081-025-00315-z

Scan to access the paper:

Motivation – wind data collected during fire experiments by CSIRO in 2015



Introduction

- Many disciplines, such as climatology, ecology, and cognitive science, collect and analyze directional/circular data.
- Linear methods may not always respect the nature of circular data.
- Wind direction data shows diverse temporal variability and modality.

Methodology

- Wrapped Autoregressive process (Breckling 1989) : **WAR**
- Direct or linked form Autoregressive model (Fisher et al. 1994): **LAR**
- Inverse Autoregressive model(Fisher et al. 1994): **IAR**
- von Mises process (Breckling 1989): **vMP**

Methodology

- Wrapped Autoregressive process (Breckling 1989) : **WAR**
- Direct or linked form Autoregressive model (Fisher et al. 1994): **LAR**
- Inverse Autoregressive model(Fisher et al. 1994): **IAR**
- von Mises process (Breckling 1989): **vMP**
- Proposed method: Hidden Markov model with conditional circular distribution:
  - **cHMM = IAR + HMM**

Existing methodology: IAR(p)

Inverse autoregressive model assumes a circular distribution with a conditional mean (Fisher et al. 1994). Let $\{\theta_t\}_{t=1}^{T}$ denote a directional time series. Then,

$$\mu_t = \mu + g\left\{\sum_{j=1}^{p} \omega_j g^{-1}(\theta_{t-j} - \mu)\right\}$$

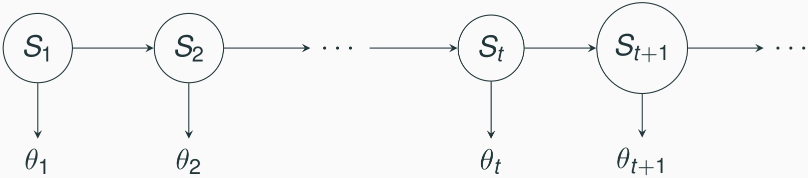
$$\theta_t|\mu_t \sim vM(\mu_t, \kappa) \Rightarrow f(\theta_t|\mu_t) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta_t - \mu_t)\},$$

where $vM(\mu, \kappa)$ denotes the von Mises distribution with mean direction $\mu \in [0, 2\pi)$ and concentration parameter $\kappa \geq 0$, and $|\omega_j| < 1; j = 1, \ldots, p$ are autocorrelation coefficients.

- Parameter vector $\mathbf{\Psi}_{IAR} = (\mu, \omega_1, \ldots, \omega_p, \kappa)^\top$.

Existing methodology: Hidden Markov models (HMMs)

Let $\{\theta_t\}_{t=1}^T$ denote a time series and $\{S_t\}_{t=1}^T$ denote the unobserved state sequence where $S_t \in \{1, 2, \ldots, K\}$.



- States follow a (first-order) Markov property: $f(S_t|S_{t-1}, S_{t-2}, \ldots) = f(S_t|S_{t-1})$
- $\{\theta_t\}_{t=1}^T$ is conditionally independent given the hidden states:

$$f(\theta_t|\theta_{t-1}, \theta_{t-2}, \ldots, S_t, S_{t-1}, \ldots) = f(\theta_t|S_t)$$

Existing methodology: Hidden Markov models (HMMs)

- Initial probabilities: $\iota_k = \mathrm{P}(S_1 = k) \in [0,1]$, and $f(\boldsymbol{\mathcal{S}}_1) = \text{Multinom}(1, \boldsymbol{\iota})$ where $\boldsymbol{\mathcal{S}}_1$ is a binary K-vector encoding S_1.
- Transition probabilities: $\gamma_{k,k'} = \mathrm{P}(S_t = k' | S_{t-1} = k) \in [0,1]$, where $\sum_{k'=1}^{K} \gamma_{k,k'} = 1$.
  We can then write the probability transition matrix

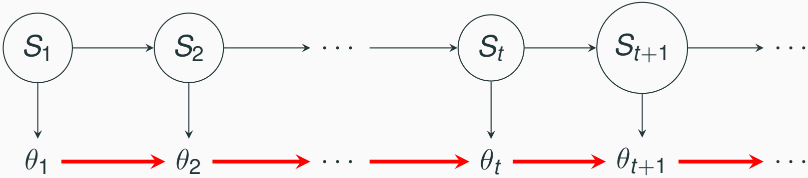
$$\boldsymbol{\Gamma} = \begin{pmatrix} \gamma_{1,1} & \gamma_{1,2} & \cdots\cdots & \gamma_{1,K} \\ \gamma_{2,1} & \gamma_{2,2} & \cdots\cdots & \gamma_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{K,1} & \gamma_{K,2} & \cdots\cdots & \gamma_{K,K} \end{pmatrix}.$$

- We have $f(\boldsymbol{\mathcal{S}}_t | S_{t-1} = k) = \text{Multinom}(1, \boldsymbol{\Gamma}_{k,\cdot})$, and hence $f(\boldsymbol{\mathcal{S}}_t | \boldsymbol{\mathcal{S}}_{t-1}) = \prod_{k=1}^{K} \prod_{k'=1}^{K} \gamma_{kk'}^{\mathcal{S}_{t-1,k}\mathcal{S}_{t,k'}}$

Proposed Method

cHMM = IAR + HMM

Circular hidden Markov model (cHMM)



In cHMMs, the directional time series $\{\theta_t\}_{t=1}^{T}$ is **not** conditionally independent given the hidden states. Specifically, we assume

$$\theta_t | S_t = k, \theta_{t-1}, \ldots, \theta_{t-p} \sim VM(\mu_{t,k}, \kappa_k); \quad \mu_{t,k} = \mu_k + g\left\{\sum_{j=1}^{p} \omega_{j,k} g^{-1}(\theta_{t-j} - \mu_k)\right\}$$

- State-specific concentration $\kappa_k > 0$ and “location” parameters $\mu_k \in [0, 2\pi)$.
- State-specific autocorrelation parameters $|\omega_{j,k}| < 1; j = 1, \ldots, p$.
- Parameter vector: $\boldsymbol{\Psi}_{cHMM} = \{\boldsymbol{\iota}^\top, \mathrm{vec}(\boldsymbol{\Gamma})^\top, \mu_1, \ldots, \mu_K, \kappa_1, \ldots, \kappa_K, \omega_{1,1}, \omega_{2,1}, \ldots, \omega_{p,K})^\top$.

cHMM: Details

1. Parameter estimation:
   - Direct numerical maximization
   - See article for marginal likelihood function and recursive formulas
2. State sequence Estimation:
   - Viterbi algorithm to determine the probable state sequence $\{\hat{S}_t\}_{t=1}^{T}$, given the estimated parameters $\hat{\mathbf{\Psi}}_{cHMM}$ (Zucchini et al. 2017)
   - See article for details

cHMM: Details

1. Parameter estimation:
   - Direct numerical maximization
   - See article for marginal likelihood function and recursive formulas
2. State sequence Estimation:
   - Viterbi algorithm to determine the probable state sequence $\{\hat{S}_t\}_{t=1}^{T}$, given the estimated parameters $\hat{\boldsymbol{\Psi}}_{cHMM}$ (Zucchini et al. 2017)
   - See article for details
3. Prediction / Forecast
   (i) Forecasting state sequence : $\{\hat{S}_{T+l|T}\}_{l=1}^{T_{new}}$.
   (ii) Forecasting direction observations given the state sequence: $\theta_{T+l|T}|\hat{S}_{T+l|T} = k \sim \mathcal{VM}(\hat{\mu}_{T+l|T,k}, \hat{\kappa}_k)$.

cHMM: Forecasting

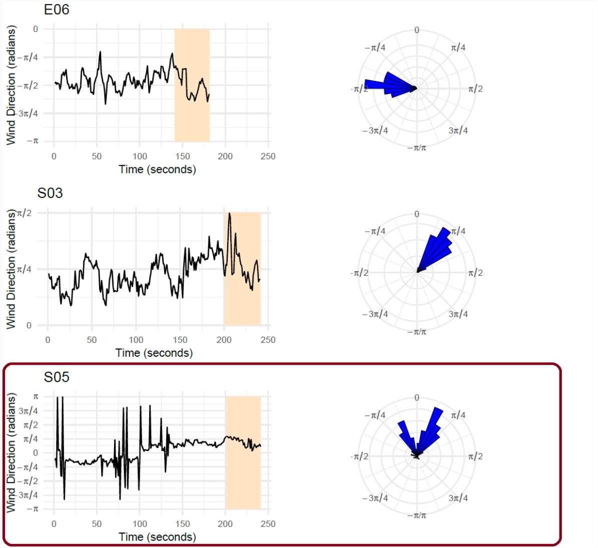
1. Forecasting state sequence $\{\hat{S}_{T+l|T}\}_{l=1}^{T_{new}}$.
   - **Unconditional forecast (UF)**: $\hat{f}(\mathcal{S}_{T+l+1|T}|\hat{S}_{T+l|T} = k) = \text{Multinom}(1, \hat{\boldsymbol{\Gamma}}_{k,\cdot})$
   - **Most probable state forecast (MPSF)**: $\hat{S}_{T+l|T} = \arg\max_{k=1,\ldots,K} \hat{P}(S_{T+l|T} = k|\{\theta_t\}_{t=1}^{T})$ (see details for how to compute this efficiently; similar to Viterbi algorithm)
   - **Last state carried forward (LSCF)**: $\hat{S}_{T+l|T} = \hat{S}_T$ for $l = 1, \ldots, T_{new}$.

cHMM: Forecasting

2. Predict/Forecast directional observation given the state sequence $\{\hat{\theta}_{T+l|T}\}_{l=1}^{T_{new}}$.
   - Simulate $\hat{\theta}_{T+l|T}|\hat{S}_{T+l|T} = k \sim vM(\hat{\mu}_{T+l|T,k}, \hat{\kappa}_k)$, where $\hat{\mu}_{T+l|T,k} = \hat{\mu}_k + g\{\sum_{j=1}^{p} \hat{\omega}_{j,k} g^{-1}(\hat{\theta}_{T+l-j|T} - \hat{\mu}_k)\}$.
   - Repeat a large number B times to construct forecast sequences $\{\{\hat{\theta}^{(b)}_{T+l|T}\}_{l=1}^{T_{new}}; b = 1, \ldots, B\}$. Point/interval/probabilistic forecasts follows from this.

Wind data application

Application – wind data collected during fire experiments by CSIRO in 2015



Application: Estimated states and in-sample performance

Based on AIC and assuming $p = 1$ autocorrelation,

- For E06, a $K = 1$ state cHMM i.e., an IAR(1) model, was chosen
- For S03, a $K = 2$ state cHMM was chosen
- For S05, a $K = 3$ state cHMM was chosen

Application: Estimated states and in-sample performance

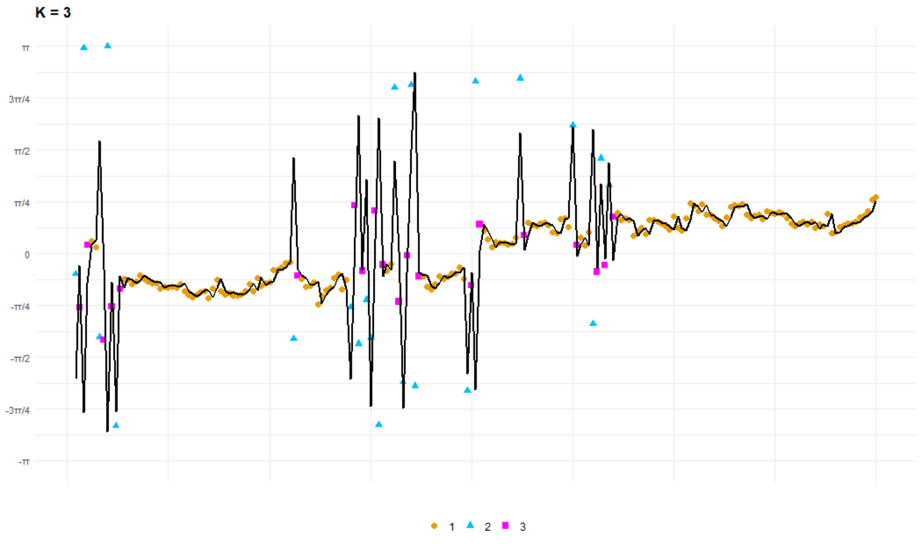
Based on AIC and assuming $p = 1$ autocorrelation,

- For E06, a $K = 1$ state cHMM i.e., an IAR(1) model, was chosen
- For S03, a $K = 2$ state cHMM was chosen
- For S05, a $K = 3$ state cHMM was chosen

In-sample performance:

Data	LAR(1)	IAR(1)	WAR(1)	vMP(1)	cHMM
			Dispersion measure (D)		
			Angular separation between observed and fitted wind directions		
E06	0.0109	0.0109	0.0109	0.0109	0.0109($k = 1$)
S03	0.0056	0.0056	0.0056	0.0056	0.0051($k = 2$)
S05	0.2256	0.2120	0.9196	0.1910	**0.0971**($k = 3$)
			MAAD		
			Circular mean absolute angular difference		
E06	0.1115	0.1116	0.1116	0.1116	0.1116($k = 1$)
S03	0.0798	0.0798	0.7970	0.0798	0.0762($k = 2$)
S05	0.5119	0.4189	1.3381	0.3727	**0.1743**($k = 3$)

Application: Estimated state sequence for S05

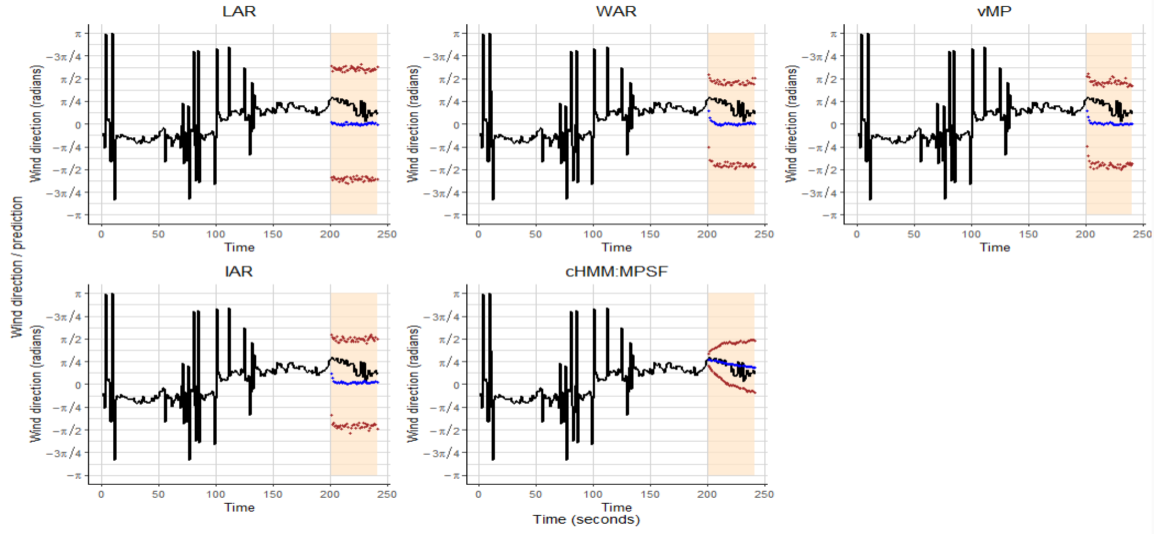


Application: Forecasting performance for S05

The interval score and circular continuous rank probability score are rules to assess performance; see article for details.

	Dispersion (D)	MAAD	Winkler/Interval score	CRPS
LAR(1)	0.193	0.591	3.804	0.198
IAR(1)	0.158	0.525	2.996	0.149
WAR(1)	0.180	0.570	2.845	0.165
vMP(1)	0.177	0.568	2.849	0.163
cHMM with UF	0.057	0.300	3.509	0.071
cHMM with MPSF	**0.022**	**0.158**	**1.374**	**0.022**
cHMM with LSCF	**0.022**	**0.158**	**1.374**	**0.022**

Application: Forecasting performance for S05



Conclusion

- cHMMs combine hidden Markov models with a conditional von Mises distribution featuring an autoregressive mean (IAR model).
- Capture time series dependency within and between states. Hidden states could be insightful e.g., wind regimes.
- Most probable state forecast (MPSF) is preferred for prediction, leveraging the Viterbi algorithm for sequential consistency.
- cHMM can be sensitive to initial values. We used a data-driven and visualization-based approach on partitioning wind directions, but more work needed.

Thank You!

Appendix: LAR(p)

$g : \Re \mapsto (-\pi, \pi)$ be an odd, monotonic and invertible function satisfying $g(0) = 0$ (Fisher et al. 1994). The linked function is $g(x) = 2\pi\{\Phi(x) - 2^{-1}\}$.

The LAR(p) model is then given by,

$$\mu_t = \beta_0 + \sum_{j=1}^{p} \beta_j g^{-1}(\theta_{t-j} - \mu)$$

$$\theta_t = \mu + g(\mu_t + e_t), \quad e_t \sim \mathrm{N}(0, \sigma^2_{LAR})$$

- Parameter vector $\boldsymbol{\Psi}_{LAR} = (\beta_0, \beta_1, \ldots, \beta_p, \mu, \sigma^2_{LAR})^\top$.

Appendix: WAR(p)

Let $\{\lambda_t\}_{t=1}^{T}$ be a latent stationary autoregressive process satisfying,

$$\lambda_t = \sum_{j=1}^{p} \beta_j \lambda_{t-j} + e_t, \qquad e_t \sim N(0, \sigma^2_{WAR})$$

The WAR(p) model is then given by (Breckling 1989),

$$\theta_t = \mu + \lambda_t \text{ (mod } 2\pi)$$

- Circular mean wind direction $\mu \in [0, 2\pi)$:

$$\mu = \arctan\left(\frac{\bar{S}}{\bar{C}}\right) + \pi I(\bar{C} < 0) + 2\pi I(\bar{C} > 0, \bar{S} < 0)$$

$$\bar{C} = T^{-1} \sum_{t=1}^{T} \cos(\theta_t) \text{ and } \bar{S} = T^{-1} \sum_{t=1}^{T} \sin(\theta_t).$$

- Parameter vector $\boldsymbol{\Psi}_{WAR} = (\mu, \beta_1, \ldots, \beta_p, \sigma^2_{WAR})^\top$.

Appendix: WAR(p)

- Centered directional time series $\{\theta_t - \hat{\mu}\}_{t=1}^{T}$.
- Sample circular autocorrelation function,

$$\hat{c}(h) = \hat{c}_{\cos}(h) + \hat{c}_{\sin}(h)$$

- Sample autocorrelation function of the $\{\lambda_t\}$,

$$\hat{C}(h) = \log\{\hat{c}(h) + 1 - \hat{c}(0)\} - \log\{1 - \hat{c}(0)\}$$

- Yule-Walker estimation is applied to the sample autocorrelation function.

$$\hat{\beta} = \hat{\boldsymbol{G}}^{-1}\hat{\boldsymbol{g}} \text{ and } \hat{\sigma}^2_{WAR} = \hat{C}(0) - \hat{\boldsymbol{g}}^T\hat{\boldsymbol{G}}^{-1}\hat{\boldsymbol{g}}$$

where the (r, s)th element of $\hat{\boldsymbol{G}}$ is given by $[\boldsymbol{G}]_{rs} = \hat{C}(r - s)$ for $r, s = 1, \ldots, p$, and $\hat{\boldsymbol{g}} = \{\hat{C}(1), \ldots, \hat{C}(p)\}^\top$.

Appendix: vM(p)

Let $\{\theta_t - \mu\}_{t=1}^{T}$ denote the centered directional time series and $(\rho_0, \ldots, \rho_p)^\top$ denote a vector of strictly positive parameters, and define the **concentration vector**,

$$\nu_{\mathbf{t}}(\boldsymbol{\rho}) = \rho_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sum_{j=1}^{p} \rho_j \begin{bmatrix} \cos(\theta_{t-j} - \mu) \\ \sin(\theta_{t-j} - \mu) \end{bmatrix}$$

$$\boldsymbol{\theta_t} | \boldsymbol{\mu_t}, \boldsymbol{\nu_{\mathrm{t}}(\rho)} \sim \boldsymbol{\mathcal{VM}(\mu_t, ||\nu_{\mathrm{t}}(\rho)||)}$$

$$\mu_t = \mathrm{Arg}\{\boldsymbol{\nu}_t(\boldsymbol{\rho})\} = \tan^{-1}\left(\frac{\nu_{t2}(\boldsymbol{\rho})}{\nu_{t1}(\boldsymbol{\rho})}\right) = \tan^{-1}\left(\frac{\sum_{j=1}^{p} \rho_j \sin(\theta_{t-j} - \mu)}{\rho_0 + \sum_{j=1}^{p} \rho_j \cos(\theta_{t-j} - \mu)}\right).$$

- parameter vector $\boldsymbol{\Psi}_{VM} = (\mu, \rho_0, \ldots, \rho_p)^\top$.

Appendix: Relationship WAR(p) and vM(p)

- Use estimates from the WAR(p) model and connect the parameters i.e. $(\beta_1, \ldots, \beta_p)^\top$ and σ^2 to vMP(p) model.

$$\sigma_j^2 = -2 \log\{ I_0(\rho_j)^{-1} I_1(\rho_j)\} \quad \text{and} \quad \sigma^2 = (\sum_{j=0}^{p} \sigma_j^{-2})^{-1} \text{(Breckling 1989).}$$

$$\sigma_j^2 = \frac{\sigma^2}{\beta_j} \quad \text{and} \quad \sigma_0^2 = \frac{\sigma^2}{1 - \sum_{j=1}^{p} \beta_j},$$

- from which it follows that,

$$\exp\left(-\frac{\sigma^2}{2\beta_j}\right) = \frac{I_1(\rho_j)}{I_0(\rho_j)} \quad \exp\left\{-\frac{\sigma^2}{2(1 - \sum_{j=1}^{p} \beta_j)}\right\} = \frac{I_1(\rho_0)}{I_0(\rho_0)}.$$

$$j = 1, \ldots, p$$

Appendix: Viterbi algorithm

$$\xi_{1,k} = \hat{\iota}_k \hat{f}(\theta_1 | S_{1,k} = 1) \qquad , \text{ k=1,...,K}$$

$$\xi_{t,k} = \max_{\{S_{t'}\}_{t'=1}^{t-1}} \hat{f}(\{S_{t'}\}_{t'=1}^{t-1}, S_t = k, \{\theta_{t'}\}_{t'=1}^{t})$$

$$\xi_{t,k} = \max_{k'} \{\xi_{(t-1),k'} \hat{\gamma}_{k',k}\} \times \hat{f}(\theta_t | S_t = k, \theta_{t-1}, \ldots, \theta_{t-p})$$

$$\boldsymbol{\hat{S}_T} = \arg \max_{k=1,\ldots,K} \boldsymbol{\xi_{T,k}} \quad \text{and} \quad \boldsymbol{\hat{S}_t} = \arg \max_{k=1,\ldots,K} \boldsymbol{\xi_{t,k} \hat{\gamma}_{k,k_{t+1}}}, \text{ t=T-1,...,1}$$

Appendix: cHMM Parameter estimation

- For a cHMM of order p, define the $K \times K$ diagonal matrix $\boldsymbol{P}(\theta_t)$,

$$\boldsymbol{P}(\theta_t) = \text{Diag}\{f(\theta_t|S_t = 1, \theta_{t-1}, \ldots, \theta_{t-p}), \ldots, f(\theta_t|S_t = K, \theta_{t-1}, \ldots, \theta_{t-p})\}$$

- The marginal likelihood function of the cHMM,

$$L_{cHMM}(\boldsymbol{\Psi}_{cHMM}) = \boldsymbol{\iota}^\top \boldsymbol{P}(\theta_1)\boldsymbol{\Gamma}\boldsymbol{P}(\theta_2)\boldsymbol{\Gamma}\ldots\boldsymbol{\Gamma}\boldsymbol{P}(\theta_T)\mathbf{1}_K$$

- Recursive computation of $L_{cHMM}(\boldsymbol{\Psi}_{cHMM})$,

$$\begin{aligned}\boldsymbol{\alpha}_1 &= \boldsymbol{P}(\theta_1)\boldsymbol{\iota}\\ \boldsymbol{\alpha}_t &= \boldsymbol{P}(\theta_t)\boldsymbol{\Gamma}^\top\boldsymbol{\alpha}_{t-1}\\ L_{cHMM}(\boldsymbol{\Psi}_{cHMM}) &= \boldsymbol{\alpha}_T^\top\mathbf{1}_K \quad \text{for } t = 2, \ldots, T\end{aligned}$$

- $\boldsymbol{\Psi}_{cHMM} = \{\boldsymbol{\iota}^\top, \text{vec}(\boldsymbol{\Gamma})^\top, \mu_1, \ldots, \mu_K, \kappa_1, \ldots, \kappa_K, \omega_{1,1}, \omega_{2,1}, \ldots, \omega_{p,K})^\top$

Appendix: Winkler score

A more explicit formula to compute the Winkler score at time point $T + l$ is given as follows:

$$W_{T+l} = \begin{cases} \text{IW}, & \text{if } \hat{L}_{T+l} < \hat{U}_{T+l} \text{ and } \theta_{T+l} \in (\hat{L}_{T+l}, \hat{U}_{T+l}) \\ \text{IW} + \frac{2}{0.05} \times \min\{\Delta(U), 2\pi - \Delta(L)\}, & \text{if } \hat{L}_{T+l} < \hat{U}_{T+l} < \theta_{T+l} \text{ and } \theta_{T+l} \notin (\hat{L}_{T+l}, \hat{U}_{T+l}) \\ \text{IW} + \frac{2}{0.05} \times \min\{\Delta(L), 2\pi - \Delta(U)\}, & \text{if } \theta_{T+l} < \hat{L}_{T+l} < \hat{U}_{T+l} \text{ and } \theta_{T+l} \notin (\hat{L}_{T+l}, \hat{U}_{T+l}) \\ 2\pi + \text{IW}, & \text{if } \hat{U}_{T+l} < \hat{L}_{T+l} \text{ and } \theta_{T+l} \in (\hat{U}_{T+l}, \hat{L}_{T+l}) \\ \hat{L}_{T+l} - \hat{U}_{T+l} + \frac{2}{0.05} \times \min\{\Delta(U), \Delta(L)\}, & \text{if } \hat{U}_{T+l} < \hat{L}_{T+l} \text{ and } \theta_{T+l} \notin (\hat{U}_{T+l}, \hat{L}_{T+l}) \end{cases},$$

where $\text{IW} = \hat{U}_{T+l} - \hat{L}_{T+l}$, $\Delta(L) = |\theta_{T+l} - \hat{L}_{T+l}|$ and $\Delta(U) = |\theta_{T+l} - \hat{U}_{T+l}|$.

Appendix: Initial value estimation for κ

$\bar{R} = \left(\bar{S}^2 + \bar{C}^2\right)^{1/2}$, where $\bar{C} = T^{-1}\sum_{t=1}^{T}\cos(\theta_t)$, $\bar{S} = T^{-1}\sum_{t=1}^{T}\sin(\theta_t)$

$$\hat{\kappa} = \begin{cases} 2\bar{R} + \bar{R}^3 + \frac{5\bar{R}^5}{6}, & \text{if } \bar{R} < 0.53 \\ -\frac{2}{5} + 1.39\bar{R} + \frac{0.43}{(1-\bar{R}}, & \text{if } 0.53 \le \bar{R} < 0.85 \\ \frac{1}{\bar{R}^3 - 4\bar{R}^2 + 3\bar{R}}, & \text{if } \bar{R} \ge 0.85 \end{cases}.$$

If $T \le 15$, we use an adjusted initial estimate $\hat{\kappa}_{adj}$ instead of $\hat{\kappa}$ which define as (Mardia et al. 2009),

$$\hat{\kappa}_{adj} = \begin{cases} \max\left\{\hat{\kappa} - \frac{2}{T\hat{\kappa}}, 0\right\}, & \text{if } \hat{\kappa} < 2 \\ \frac{(T-1)^3\hat{\kappa}}{T^3 + T}, & \text{if } \hat{\kappa} \ge 2. \end{cases}$$

Appendix: Initial value estimation for ω_j

Expression for the sample autocorrelation given by (Fisher et al. 1994),

$$\hat{\omega}_j = \frac{\det(\sum_{t=1}^{T-j} X_t X_{t+j}^T)}{\{\det(\sum_{t=1}^{T-j} X_t X_t^T) \det(\sum_{t=j+1}^{T} X_t X_t^T)\}^{1/2}}; \; j = 1, \ldots, p,$$

where $X_t = \{\cos(\theta_t), \sin(\theta_t\}^\top$ and $\det(\cdot)$ denotes the determinant operator.

Appendix: In-sample performance

- **Dispersion** : A measure based on the angular separation between the observed and fitted wind directions (Mardia et al. 2009; Harvey et al. 2024),

$$D = (T-1)^{-1} \sum_{t=2}^{T} \{1 - \cos(\theta_t - \hat{\theta}_t)\}$$

- **Mean absolute angular difference (MAAD)** : circular mean of absolute angular difference $\zeta(\theta_t, \hat{\theta}_t)$,

$$\zeta(\theta_t, \hat{\theta}_t) = \begin{cases} |\theta_t - \hat{\theta}_t| & \text{if } |\theta_t - \hat{\theta}_t| \leq \pi \\ 2\pi - |\theta_t - \hat{\theta}_t| & \text{if } \pi < |\theta_t - \hat{\theta}_t| \leq 2\pi \end{cases}$$

Appendix: Forecasting performance

- **Point forecast performance:**
  - Dispersion
  - MAAD

- **Interval forecast performance:**
  - Circular average of Winkler score (interval score):

$$\mathrm{W}_t = \text{Interval width} + 2 \times \alpha^{-1} \times \mathbb{I}\{\theta_t \notin (\hat{L}_{T+l}, \hat{U}_{T+l})\} \times \Delta_t$$

  Δ_t denotes the distance between θ_t and the closest end point of the $(1-\alpha)\%$ interval.

- **Probabilistic forecast performance:**
  - Continuous ranked probability score (CRPS) (Grimit et al. 2006).

$$\mathrm{CRPS}(\theta_t) = \mathbb{E}\left\{\zeta(\Theta, \theta_t)\right\} - 2^{-1}\mathbb{E}\left\{\zeta(\Theta, \Theta^*)\right\}$$

  Θ and Θ^* denote independent copies of directional random variables.
  `CircSpaceTime` package (Lasinio et al. 2020) used to compute the circular CRPS.