GEYSERS, WIND, FINANCIAL RETURNS AND HOMICIDES; APPLICATIONS OF HIDDEN MARKOV MODELS

Walter Zucchini
University of Göttingen

Iain MacDonald
University of Cape Town

Applications of Hidden Markov Models

1. Eruptions of the Old Faithful geyser
2. Wind direction at Koeberg
3. Daily returns on the Tokyo Stock Price Index (TOPIX)
4. Cape Town homicides and suicides

Eruptions of Old Faithful geyser

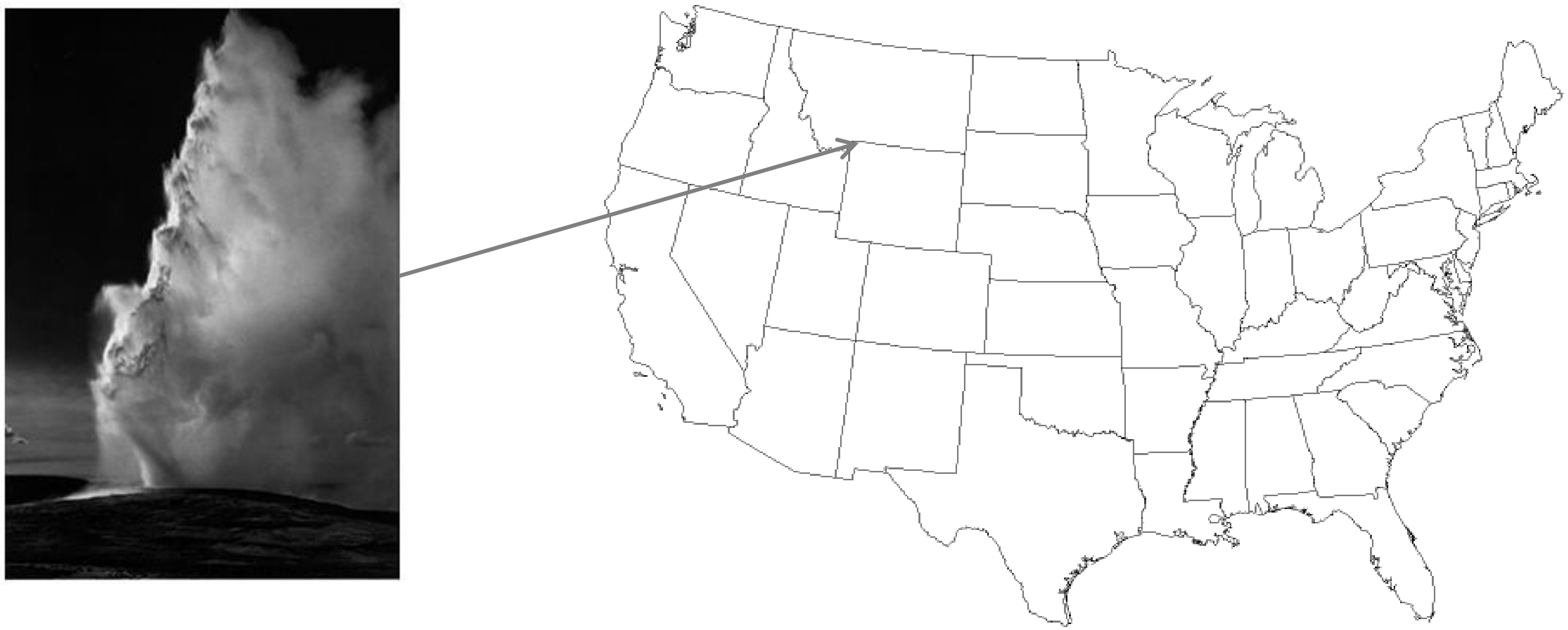
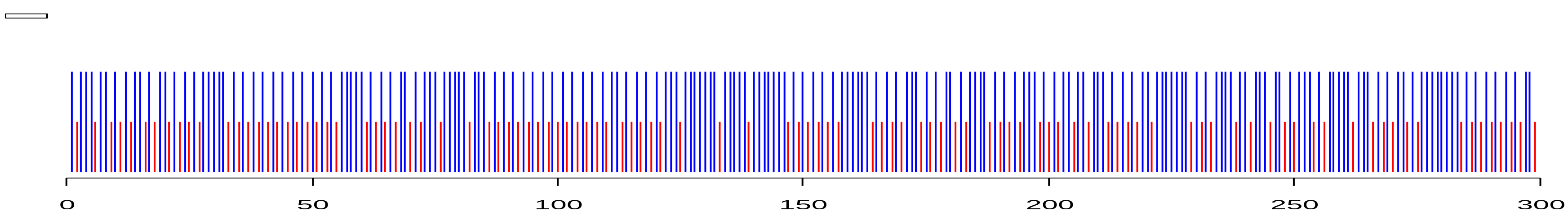


Photo by Ansel Adams

Data: Successive eruptions 01.08.1985 – 15.08.1985 (299 observations)
Classified as **short** ($x_t = 0$) or **long** ($x_t = 1$)



Series $X_1 \longrightarrow X_2 \longrightarrow X_3 \longrightarrow X_4 \longrightarrow$

Notation: $X^{(t)}$ denotes the history up to time t, i.e. $\{X_t, X_{t-1}, \cdots, X_1\}$.

Markov property: $\Pr(X_t \,|\, X^{(t-1)}) \;=\; \Pr(X_t \,|\, X_{t-1})$

Transition probability matrix of the homogeneous Markov chain:

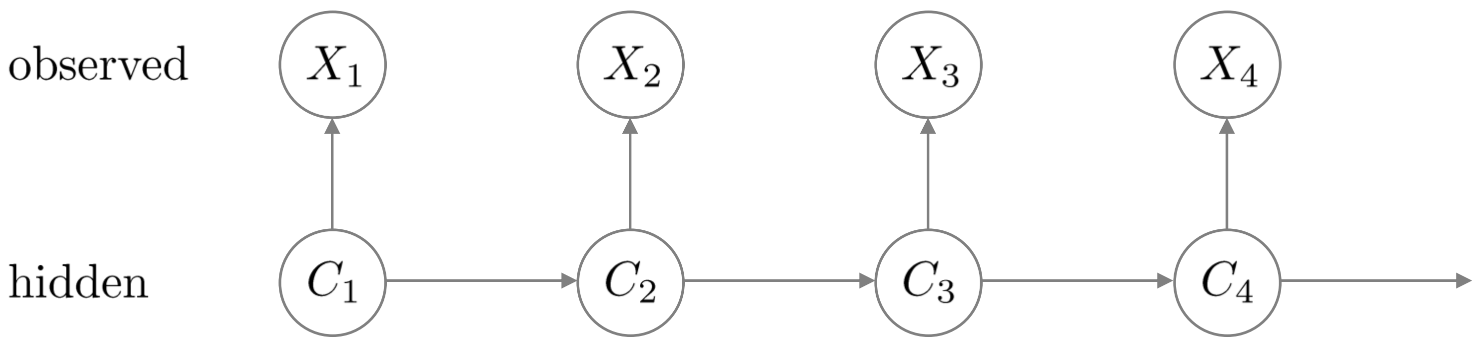
$$\Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} \Pr(0 \rightarrow 0) & \Pr(0 \rightarrow 1) \\ \Pr(1 \rightarrow 0) & \Pr(1 \rightarrow 1) \end{pmatrix} \qquad \begin{array}{l} \gamma_{11} + \gamma_{12} = 1 \\ \gamma_{21} + \gamma_{22} = 1 \end{array}$$

Initial state distribution: $\delta = (\delta_1 \;\; \delta_2)$

If the chain is also stationary: $\delta = \delta\,\Gamma$

$$\widehat{\Gamma} = \begin{pmatrix} 0.00 & 1.00 \\ 0.54 & 0.46 \end{pmatrix} \qquad \widehat{\delta}' = \begin{pmatrix} 0.35 \\ 0.65 \end{pmatrix}$$

State series:	$C_1, C_2, \cdots$	homogeneous two-state Markov chain
Observed series:	$X_1, X_2, \cdots$	mixture of two Bernoulli distributions
Assumption:		conditional independence



Definition of a HMM

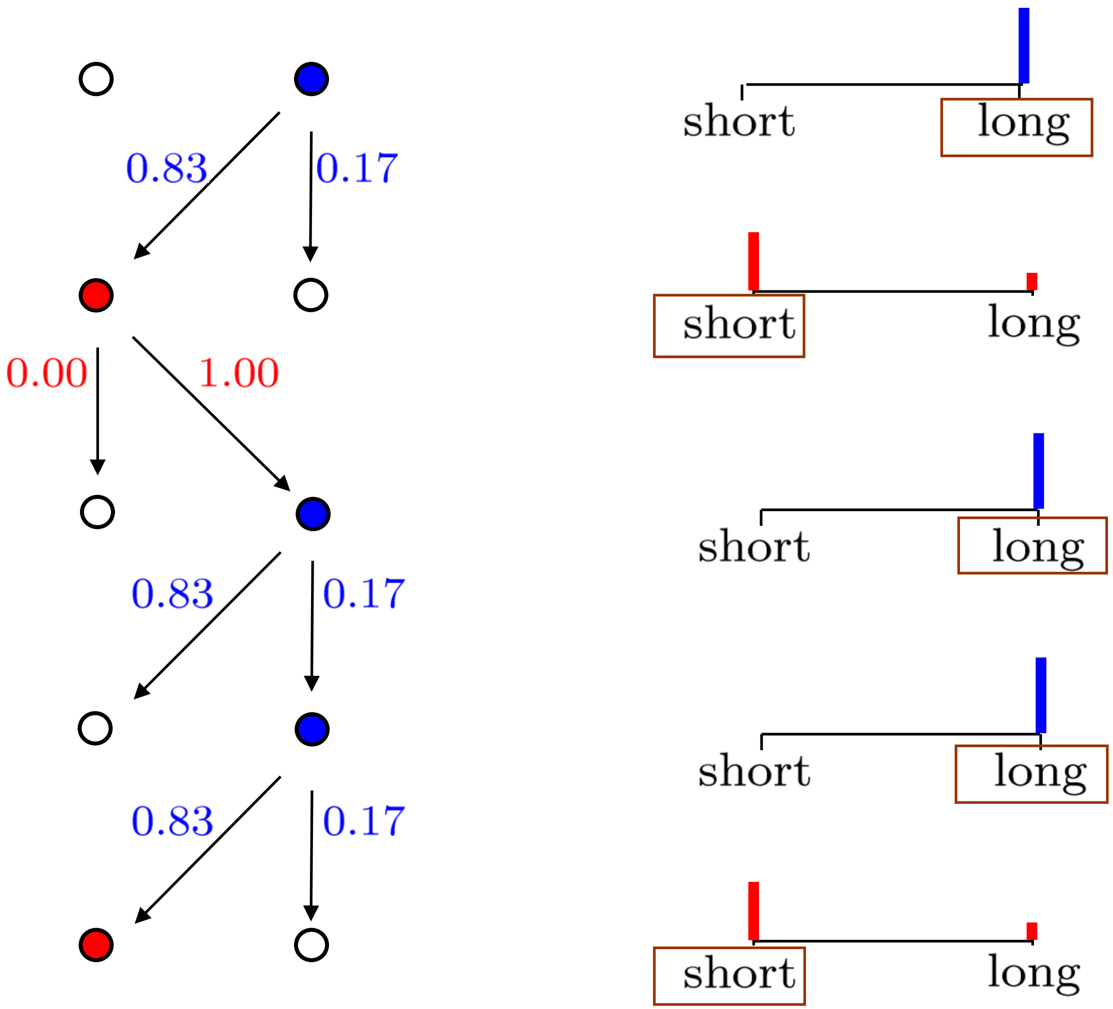
Markov property $\Pr(C_t \,|\, C^{(t-1)}) = \Pr(C_t \,|\, C_{t-1})$

Conditional independence $\Pr(X_t \,|\, X^{(t-1)}, C^{(t)}) = \Pr(X_t \,|\, C_t)$

State-dependent distributions $\begin{cases} X_t \,|\, C_t = 1 & \sim \quad \text{Bernoulli}(\pi_1) \\ X_t \,|\, C_t = 2 & \sim \quad \text{Bernoulli}(\pi_2) \end{cases}$

Two-state Bernoulli HMM

unobserved state		state-dependent distribution	transition prob. matrix
state 1 0.35	state 2 0.65	$\pi_1 = P(\text{long} \mid \text{state 1}) = 0.23$ $\pi_2 = P(\text{long} \mid \text{state 2}) = 1.00$	$\Gamma = \begin{pmatrix} 0.00 & 1.00 \\ 0.83 & 0.17 \end{pmatrix}$



observation

long ($x_1 = 1$)

short ($x_2 = 0$)

long ($x_3 = 1$)

long ($x_4 = 1$)

short ($x_5 = 0$)

hidden

<u>observation</u>

long ($x_1 = 1$)

short ($x_2 = 0$)

long ($x_3 = 1$)

long ($x_4 = 1$)

short ($x_5 = 0$)

Likelihood of an homogeneous HMM

$$\boxed{L_T = \delta\, \mathrm{P}(x_1)\, \Gamma\, \mathrm{P}(x_2)\, \Gamma\, \mathrm{P}(x_3) \cdots\ \Gamma\, \mathrm{P}(x_T)\, 1'}$$

Two-state HMM

$$\Gamma\, \mathrm{P}(x_t) = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} p_1(x_t) & 0 \\ 0 & p_2(x_t) \end{pmatrix}$$

Three-state HMM

$$\Gamma\, \mathrm{P}(x_t) = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{pmatrix} \begin{pmatrix} p_1(x_t) & 0 & 0 \\ 0 & p_2(x_t) & 0 \\ 0 & 0 & p_3(x_t) \end{pmatrix}$$

State-dependent distributions: $p_1(x), p_2(x), p_3(x)$.

Likelihood of a two-state Bernoulli-HMM

Observations: $x_1 \quad x_2 \quad x_3 \quad \cdots \quad x_T$

Likelihood: $\delta\,\mathrm{P}(x_1) \quad \Gamma\,\mathrm{P}(x_2) \quad \Gamma\,\mathrm{P}(x_3) \quad \cdots \quad \Gamma\,\mathrm{P}(x_T) \quad 1'$

Two-state Bernoulli-HMM:

$$\Gamma\mathrm{P}(x) = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} {\pi_1}^x(1-\pi_1)^{1-x} & 0 \\ 0 & {\pi_2}^x(1-\pi_2)^{1-x} \end{pmatrix}$$

Model parameters: γ_{12}, γ_{21}, π_1, π_2 (and δ_1 if non-stationary)

Parameter estimation:
– EM algorithm (Baum-Welch algorithm), or
– direct numerical maximization (e.g. `nlm` in `R`).

Observations: $x_1 \quad x_2 \quad x_3 \quad \cdots \quad x_T$

Likelihood: $\delta\,\mathrm{P}(x_1) \quad \Gamma\,\mathrm{P}(x_2) \quad \Gamma\,\mathrm{P}(x_3) \quad \cdots \quad \Gamma\,\mathrm{P}(x_T) \quad 1'$

Three-state Poisson-HMM

$$\mathbf{\Gamma P}(x) = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{pmatrix} \begin{pmatrix} \frac{e^{-\lambda_1}\lambda_1^x}{x!} & 0 & 0 \\ 0 & \frac{e^{-\lambda_2}\lambda_2^x}{x!} & 0 \\ 0 & 0 & \frac{e^{-\lambda_3}\lambda_3^x}{x!} \end{pmatrix}$$

Model parameters: $\begin{pmatrix} - & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & - & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & - \end{pmatrix}, \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3. \end{pmatrix}$

Azzalini and Bowman (1990) fitted

- a first order Markov chain: The fit is reasonable, except the acf.
- a second order Markov chain: The fit is reasonable, including the acf.

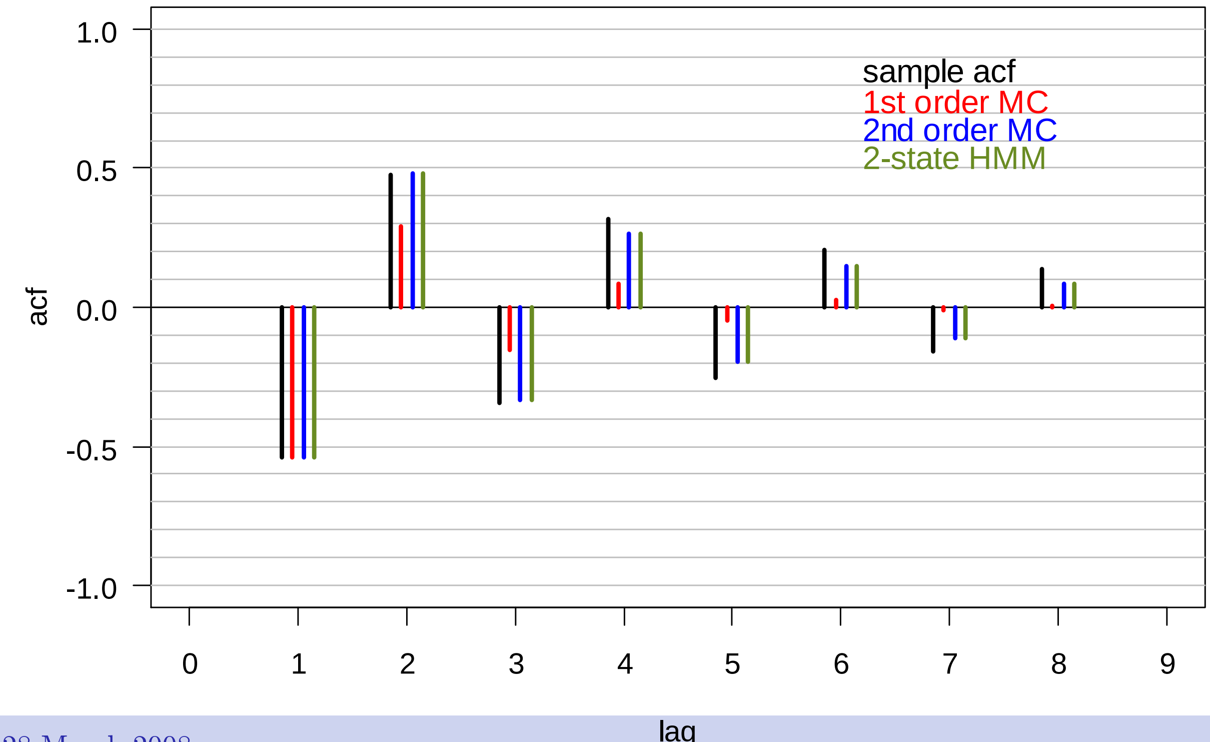
MacDonald and Zucchini (1997) fitted

- a two-state Binary-HMM. The fit is reasonable, including the acf.

Estimates

$$\widehat{\Gamma} = \begin{pmatrix} 0.00 & 1.00 \\ 0.83 & 0.17 \end{pmatrix}$$

$$\widehat{\pi} = \begin{pmatrix} 0.23 & 1.00 \end{pmatrix}$$



Additional issues:
– model checking - pseudo residuals
– model selection
– forecasting and monitoring
– decoding (identify the most likely states)

Model selection criteria

model	k	$-\,\mathrm{L}_T$	AIC	BIC
1-state hidden Markov (indep.)	1	193.80	389.60	393.31
First-order Markov chain	2	134.24	272.48	279.88
Second-order Markov chain	4	127.12	262.24	277.04
2-state hidden Markov	4	127.31	262.62	277.42
3-state hidden Markov	9	126.85	271.70	305.00
4-state hidden Markov	16	126.59	285.18	344.39
2-state second-order HM	6	126.90	265.80	288.00

Interval-censored durations

t	recorded
1	4.0
2	2.1
3	long
4	long
5	long
6	short
7	4.4
8	4.3
.	.
241	1.9
242	4.4
243	medium
244	long
245	2.0
246	long
247	3.3
248	1.8
.	.
295	4.1
296	2.1
297	long
298	long
299	short

recorded	interval (min)
short	$0.0 \leq \text{duration} \leq 3.0$
medium	$2.5 \leq \text{duration} \leq 3.5$
long	$3.0 \leq \text{duration}$

$$\text{obs:} \quad x_1^-, x_1^+ \quad x_2^-, x_2^+ \quad x_3^-, x_3^+ \quad \cdots \quad x_T^-, x_T^+$$

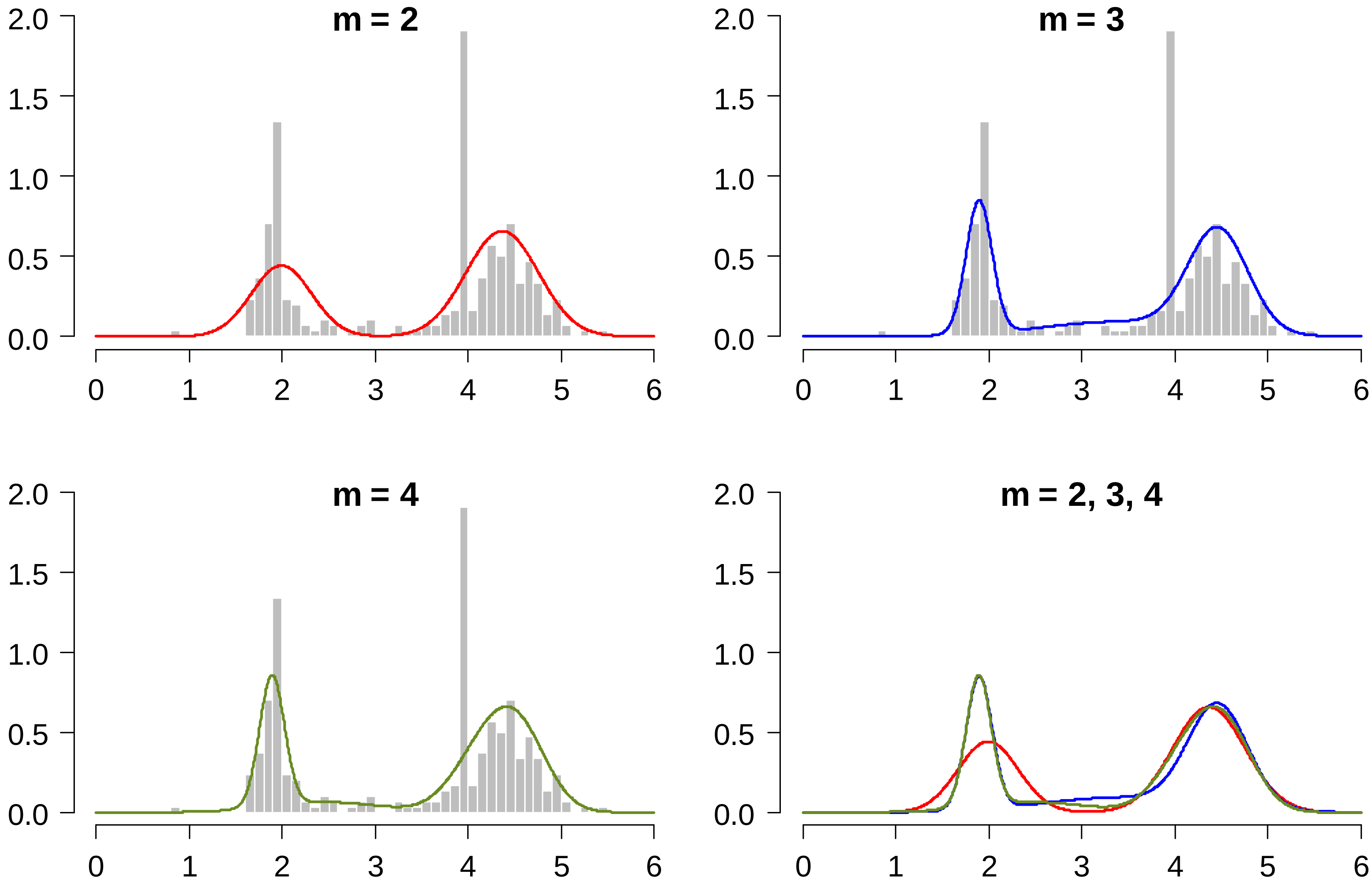
$$L_T = \quad \delta\, \mathrm{P}(x_1) \quad \Gamma\, \mathrm{P}(x_2) \quad \Gamma\, \mathrm{P}(x_3) \quad \cdots \quad \Gamma\, \mathrm{P}(x_T) \quad 1'$$

$$\Gamma\ \mathrm{P}(x_t) = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{pmatrix} \begin{pmatrix} p_1(x_t) & 0 & 0 \\ 0 & p_2(x_t) & 0 \\ 0 & 0 & p_3(x_t) \end{pmatrix}$$

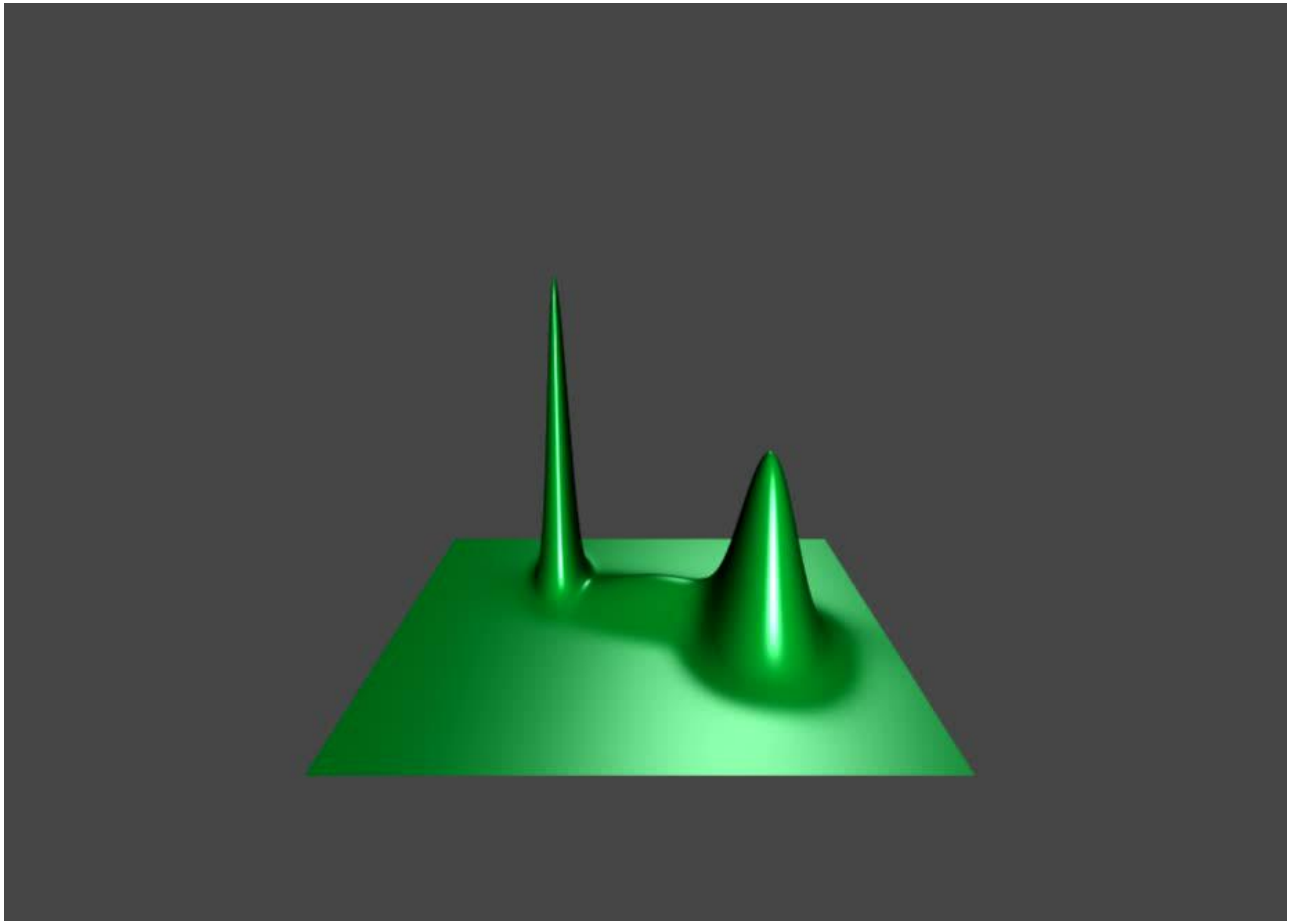
$$p_i(x_t) = \int_{x_t^-}^{x_t^+} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(z-\mu_i)^2}{2\sigma_i{}^2}} dz$$

$$= \Phi\left(\frac{x_t^+ - \mu_i}{\sigma_i}\right) - \Phi\left(\frac{x_t^- - \mu_i}{\sigma_i}\right)$$

Marginal distribution of duration



$$\text{Given state } i: \quad (x_t, y_t) \sim \text{bivariate normal}, \quad i = 1, 2, 3$$



Wind direction at Koeberg

Data:	Average hourly wind direction and speed
Period:	01.05.1985 – 30.04.1989
Length:	35 064 observations

Aim: Short-term forecasting for radioactive plume modelling

Models for the hourly wind direction

0. First-order Markov chain — baseline model
1. Multinomial–HMMs
2. Two-state seasonal multinomial–HMM

Observations: One of 16 compass directions
Code: N=1, NNE=2, ..., NNW=16

Models for the hourly change in wind direction

1. Von-Mises-HMM
2. Von-Mises-HMM with wind speed as covariate

Observations: Wind speed (s) and direction (x)

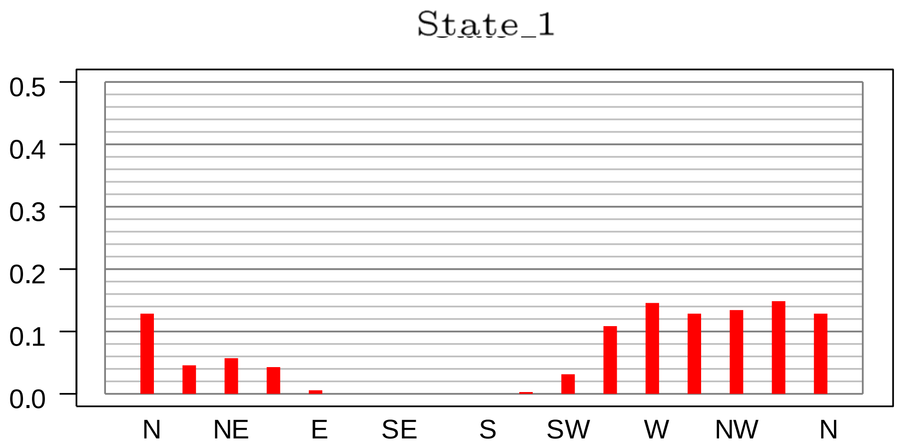
State–dependent model: $\Pr(X_t = j \mid C_t = i) = \begin{cases} \pi_{j1}, & \text{for } i = 1 \\ \pi_{j2}, & \text{for } i = 2 \end{cases}$

Estimates

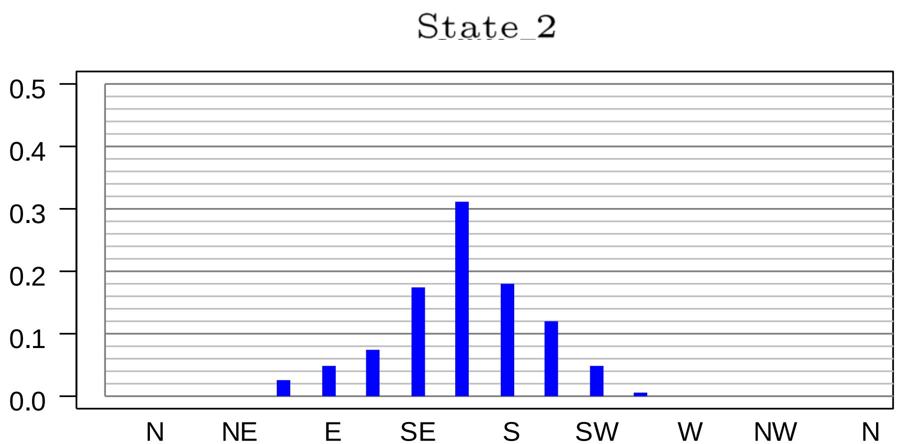
$$\hat{\Gamma} = \begin{pmatrix} 0.964 & 0.036 \\ 0.031 & 0.969 \end{pmatrix}$$

$$\hat{\delta}' = \begin{pmatrix} 0.462 \\ 0.538 \end{pmatrix}$$

j	Direction	π_{j1}	π_{j2}
1	N	0.129	0.000
2	NNE	0.048	0.000
3	NE	0.059	0.001
4	ENE	0.044	0.026
5	E	0.006	0.050
6	ESE	0.001	0.075
7	SE	0.000	0.177
8	SSE	0.000	0.313
9	S	0.001	0.181
10	SSW	0.004	0.122
11	SW	0.034	0.048
12	WSW	0.110	0.008
13	W	0.147	0.000
14	WNW	0.130	0.000
15	NW	0.137	0.000
16	NNW	0.149	0.000



"North–west"
(Most probable: NNW)



"South–east"
(Most probable: SSE)

State–dependent model: $\Pr(X_t = j \mid C_t = i) = \begin{cases} \pi_{j1}, & \text{for } i = 1 \\ \pi_{j2}, & \text{for } i = 2 \\ \pi_{j3}, & \text{for } i = 3 \end{cases}$

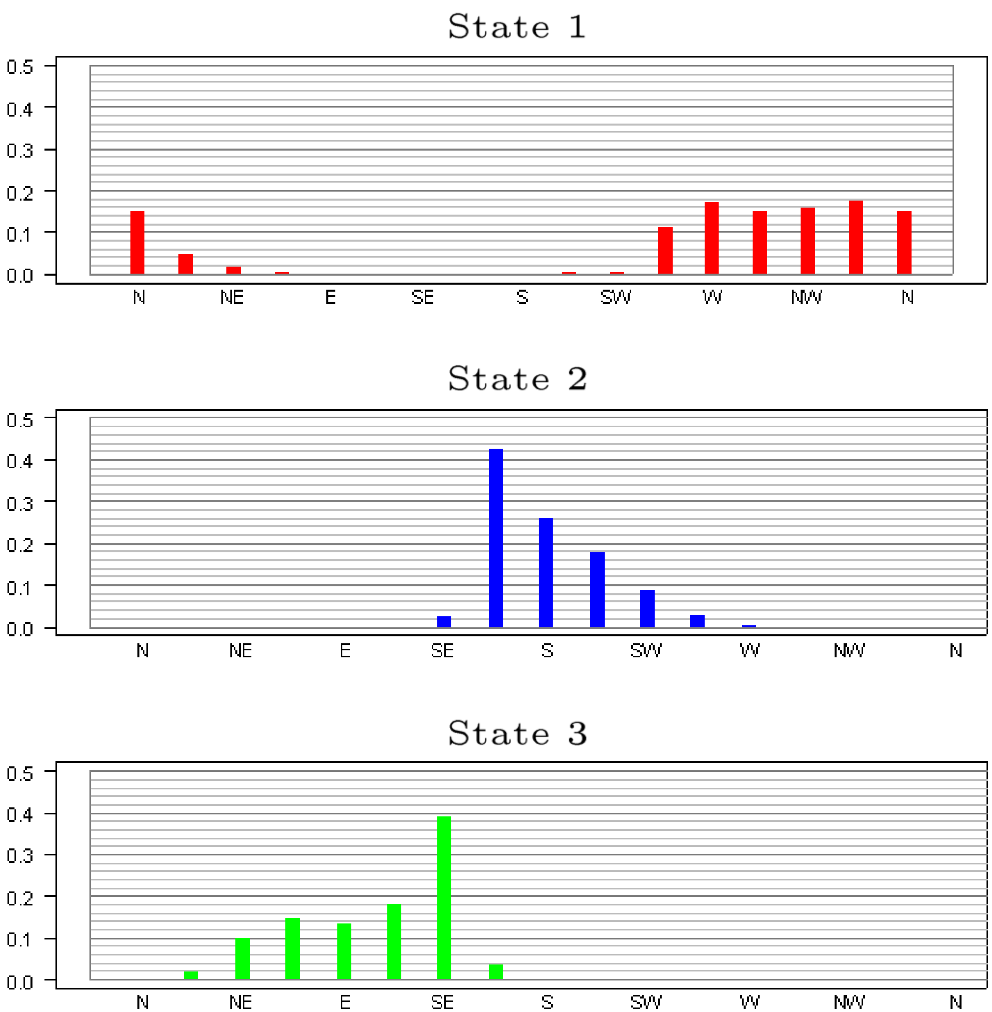
Estimates

$$\hat{\Gamma} = \begin{pmatrix} 0.957 & 0.030 & 0.013 \\ 0.015 & 0.923 & 0.062 \\ 0.051 & 0.077 & 0.872 \end{pmatrix}$$

$$\hat{\delta}' = \begin{pmatrix} 0.400 \\ 0.377 \\ 0.223 \end{pmatrix}$$

j	Direction	π_{j1}	π_{j2}	π_{j3}
1	N	0.148	0.000	0.001
2	NNE	0.047	0.000	0.016
3	NE	0.016	0.000	0.097
4	ENE	0.003	0.000	0.148
5	E	0.001	0.000	0.132
6	ESE	0.000	0.000	0.182
7	SE	0.000	0.023	0.388
8	SSE	0.000	0.426	0.033
9	S	0.000	0.257	0.002
10	SSW	0.002	0.176	0.000
11	SW	0.020	0.089	0.000
12	WSW	0.111	0.028	0.000
13	W	0.169	0.002	0.000
14	WNW	0.151	0.000	0.000
15	NW	0.159	0.000	0.001
16	NNW	0.173	0.000	0.000

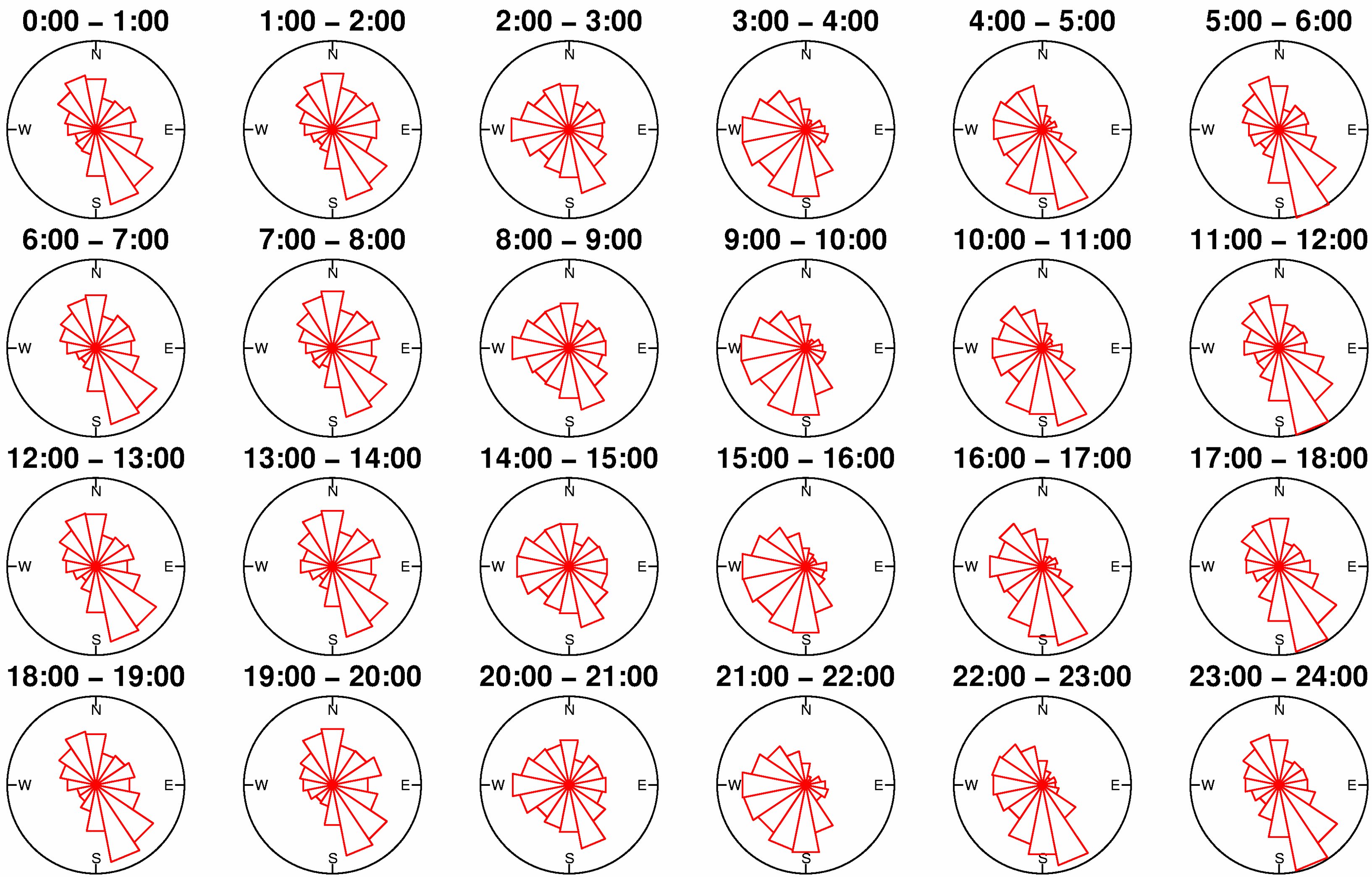
Three-state multinomial-HMM – Estimates of state-dependent distributions



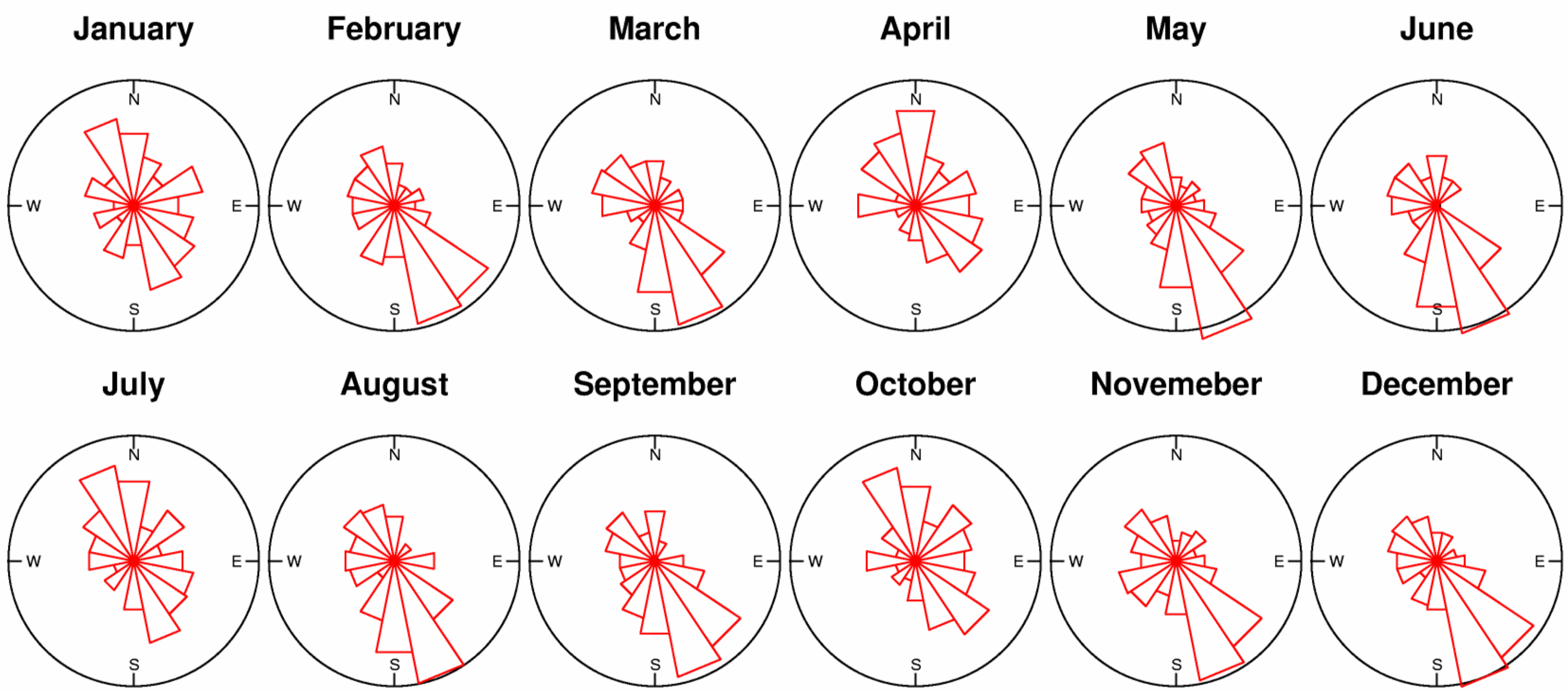
"North–west"
(Most probable: NNW)

"South–east"
is split in two

Wind direction at Koeberg by time of day



Wind direction at Koeberg (23:00 – 24:00)



Two-state seasonal multinomial-HMM

State–dependent model $\Pr(X_t = j \mid C_t = i) = \begin{cases} \pi_{j1}, & \text{for } i = 1 \\ \pi_{j2}, & \text{for } i = 2 \end{cases}$

$$\Gamma = \begin{pmatrix} \gamma_{11}(t) & \gamma_{12}(t) \\ \gamma_{21}(t) & \gamma_{22}(t) \end{pmatrix}$$

Transition probabilities are now functions of a covariate, **time**.

$$\text{logit}(\gamma_{12}(t)) = a_1 + b_1 \cos\left(\frac{2\pi t}{24}\right) + c_1 \sin\left(\frac{2\pi t}{24}\right) + d_1 \cos\left(\frac{2\pi t}{8766}\right) + e_1 \sin\left(\frac{2\pi t}{8766}\right)$$

$$\text{logit}(\gamma_{21}(t)) = a_2 + \underbrace{b_2 \cos\left(\frac{2\pi t}{24}\right) + c_2 \sin\left(\frac{2\pi t}{24}\right)}_{\text{daily cycle}} + \underbrace{d_2 \cos\left(\frac{2\pi t}{8766}\right) + e_2 \sin\left(\frac{2\pi t}{8766}\right)}_{\text{annual cycle}}$$

Two-state seasonal multinomial-HMM - estimates

State–dependent model $\Pr(X_t = j \mid C_t = i) = \begin{cases} \pi_{j1}, & \text{for } i = 1 \\ \pi_{j2}, & \text{for } i = 2 \end{cases}$

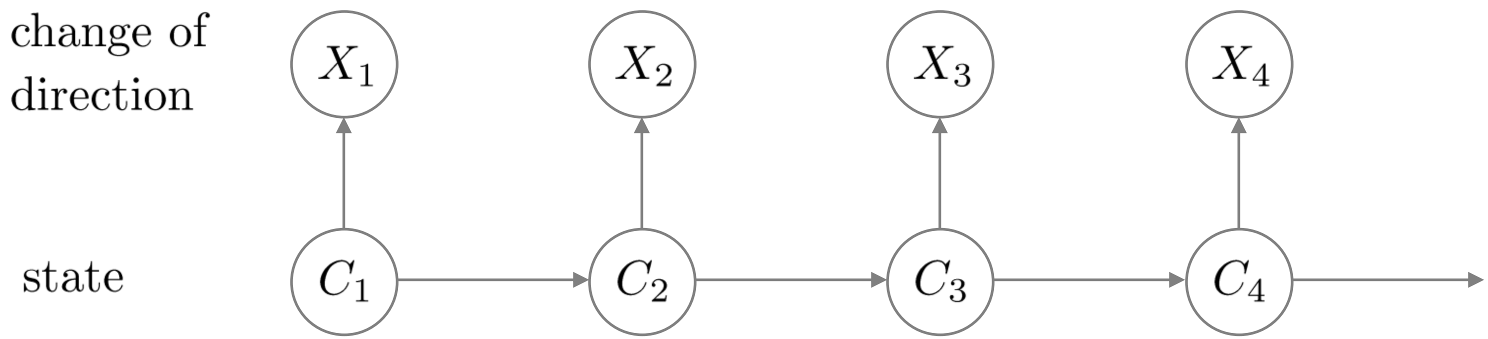
Parameters of $\Gamma(t)$

	$i = 1$	$i = 2$
$\hat{a}_i$	-3.349	-3.523
$\hat{b}_i$	0.197	-0.272
$\hat{c}_i$	-0.695	0.801
$\hat{d}_i$	-0.208	0.082
$\hat{e}_i$	-0.401	-0.089

This model has fewer parameters (40) than the three-state HMM (51), which doesn't take seasonality into account.

j	Direction	π_{j1}	π_{j2}
1	N	0.127	0.000
2	NNE	0.047	0.000
3	NE	0.057	0.002
4	ENE	0.027	0.040
5	E	0.004	0.052
6	ESE	0.001	0.076
7	SE	0.001	0.179
8	SSE	0.000	0.317
9	S	0.001	0.183
10	SSW	0.007	0.121
11	SW	0.059	0.026
12	WSW	0.114	0.003
13	W	0.145	0.000
14	WNW	0.128	0.000
15	NW	0.135	0.000
16	NNW	0.147	0.000

Model 1 for change in direction: von Mises-HMM



Likelihood: $L_T = \delta \, \mathrm{P}(x_1) \, \Gamma \, \mathrm{P}(x_2) \, \Gamma \, \mathrm{P}(x_3) \cdots \, \Gamma \, \mathrm{P}(x_T) \, 1'$

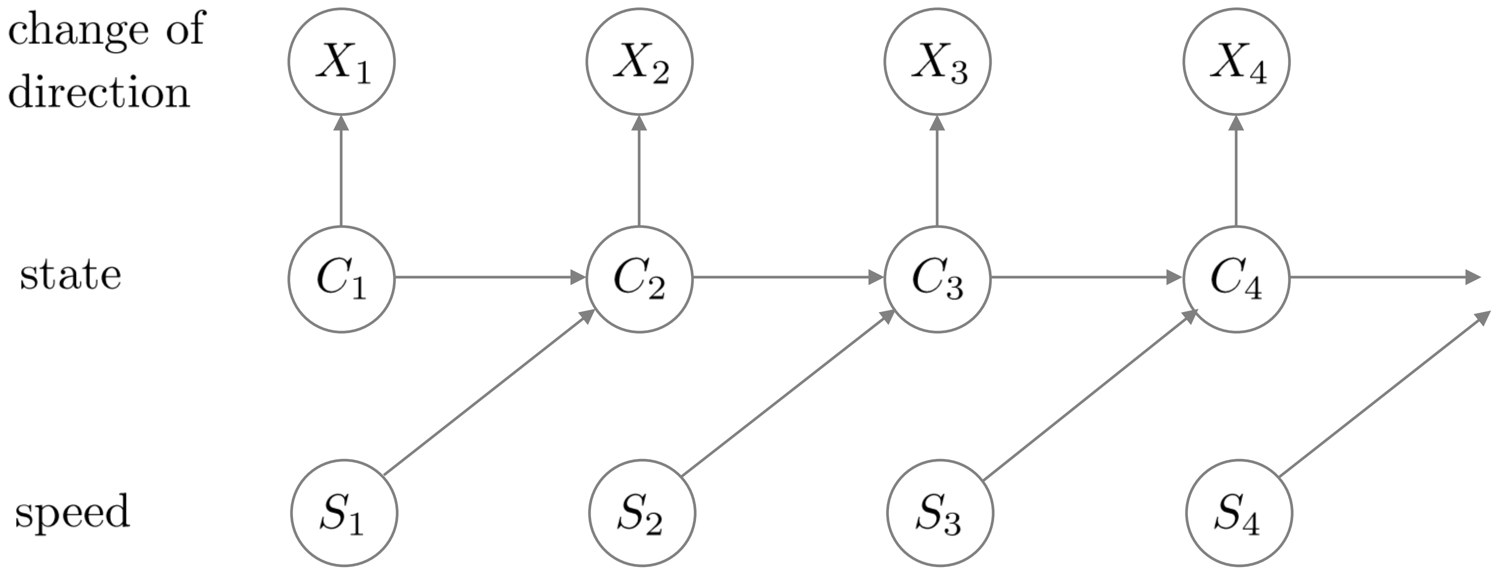
$$\Gamma \mathrm{P}(x_t) = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} p_1(x_t) & 0 \\ 0 & p_2(x_t) \end{pmatrix}$$

Von Mises distribution $\mathrm{vM}(\mu_i, \kappa_i)$:

$$p_i(x) = \frac{1}{2\pi I_0(\kappa_i)} \, e^{\kappa_i \cos(x - \mu_i)}, \quad \kappa_i \geq 0, \; \mu_i \in [0, 2\pi),$$

where $I_0(\kappa)$ is the modified Bessel function of order 0.

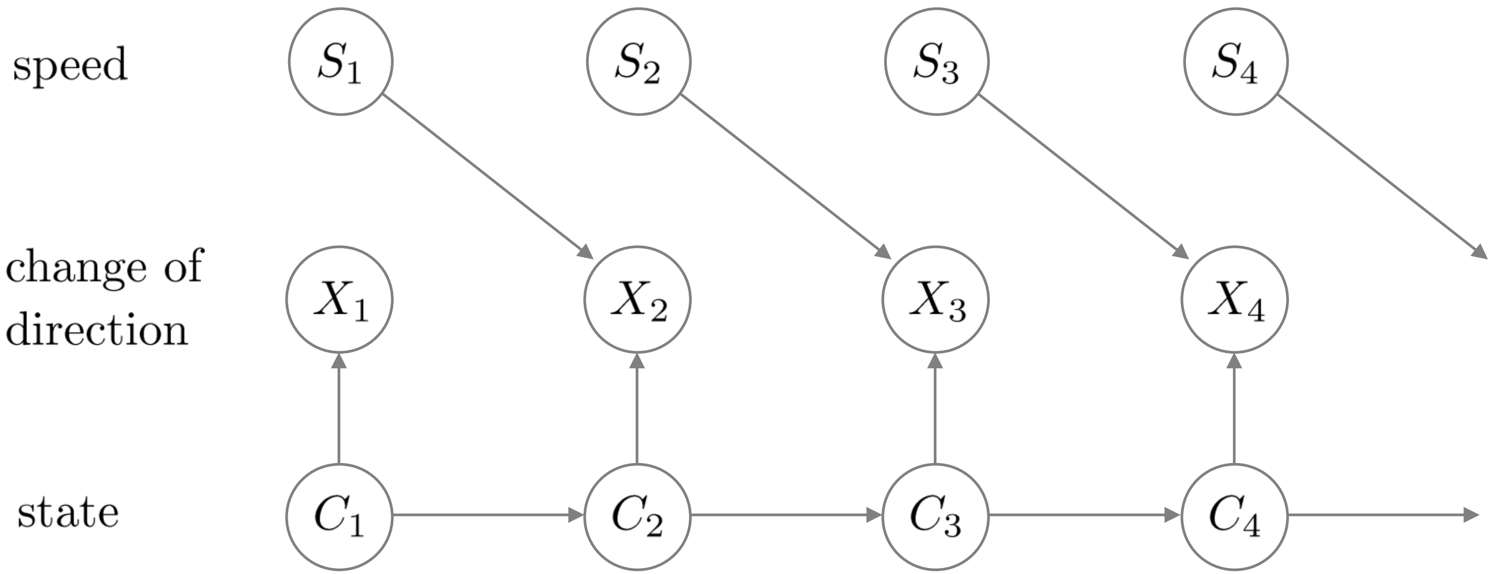
Model 2 for change in direction: von Mises-HMM – wind speed affects Γ



Transition probability matrix is a function of of s_{t-1}, e.g.,

(a) $\Gamma(t) = g(s_{t-1})$
(b) $\Gamma(t) = g(\sqrt{s_{t-1}})$

Model 3 for change in direction: von Mises-HMM – wind speed affects κ



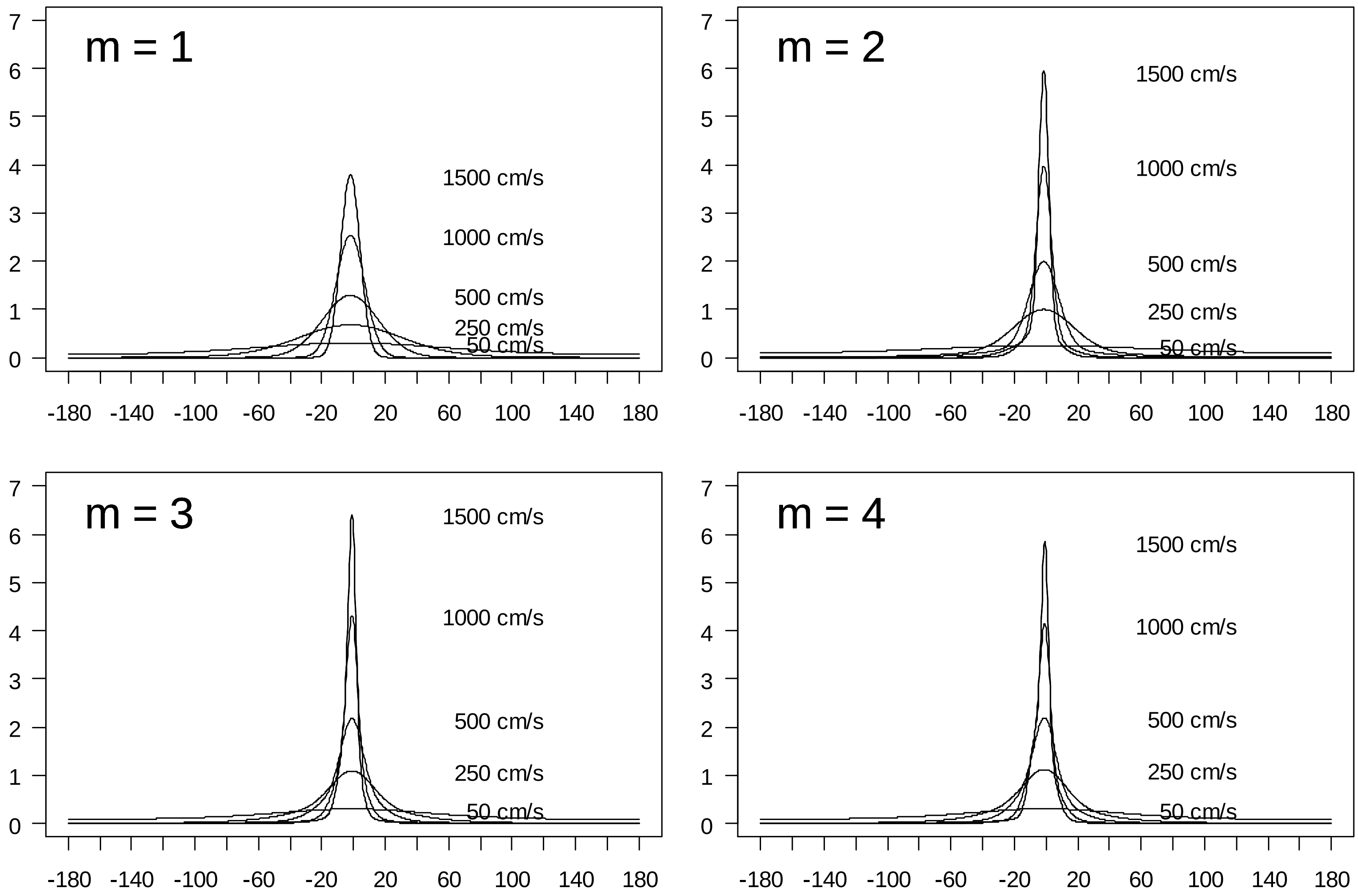
Idea: The higher the wind speed the less likely it will change direction.

Von Mises parameter κ is a function of s_{t-1}, e.g.,

(a) $\log \kappa_i = \zeta_{0i} + \zeta_{1i}\sqrt{s_{t-1}}$

(b) $\kappa_i = \zeta_{0i} + \zeta_{1i} s_{t-1}^2$ $\Leftarrow$ the best we could find

Model 3(b) – marginal distributions for increasing numbers of states



Model 3(b) estimates – four-state Von-Mises-HMM

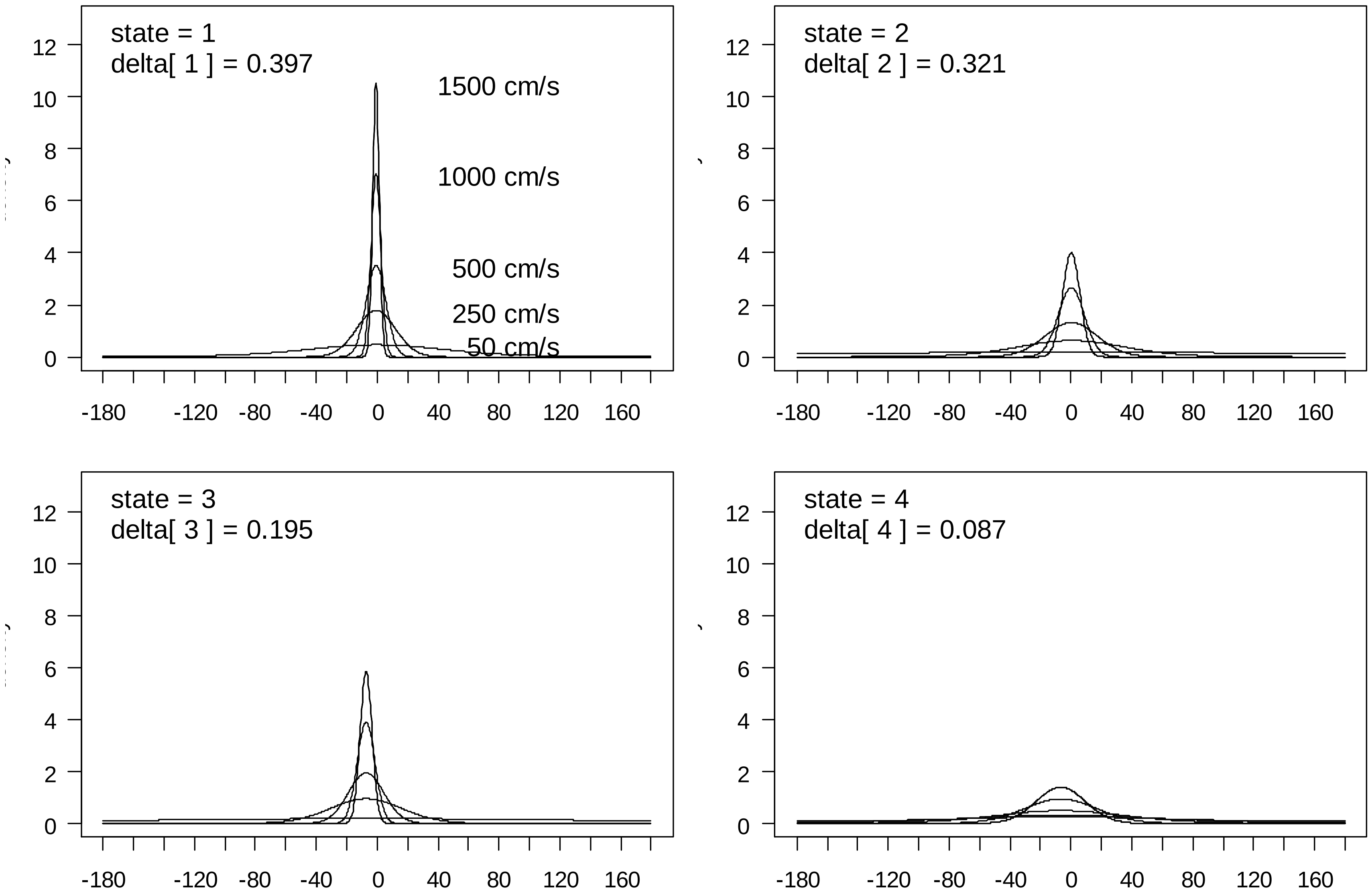
$$\widehat{\Gamma} = \begin{pmatrix} 0.755 & 0.163 & 0.080 & 0.003 \\ 0.182 & 0.707 & 0.045 & 0.006 \\ 0.185 & 0.000 & 0.722 & 0.093 \\ 0.031 & 0.341 & 0.095 & 0.533 \end{pmatrix}$$

$$p_i(x) = \frac{1}{2\pi I_0(\kappa_i)} \; e^{\kappa_i \cos(x-\mu_i)},$$

$$\kappa_i = \zeta_{0i} + \zeta_{1i} s_{t-1}^2, \quad i = 1, 2, 3, 4$$

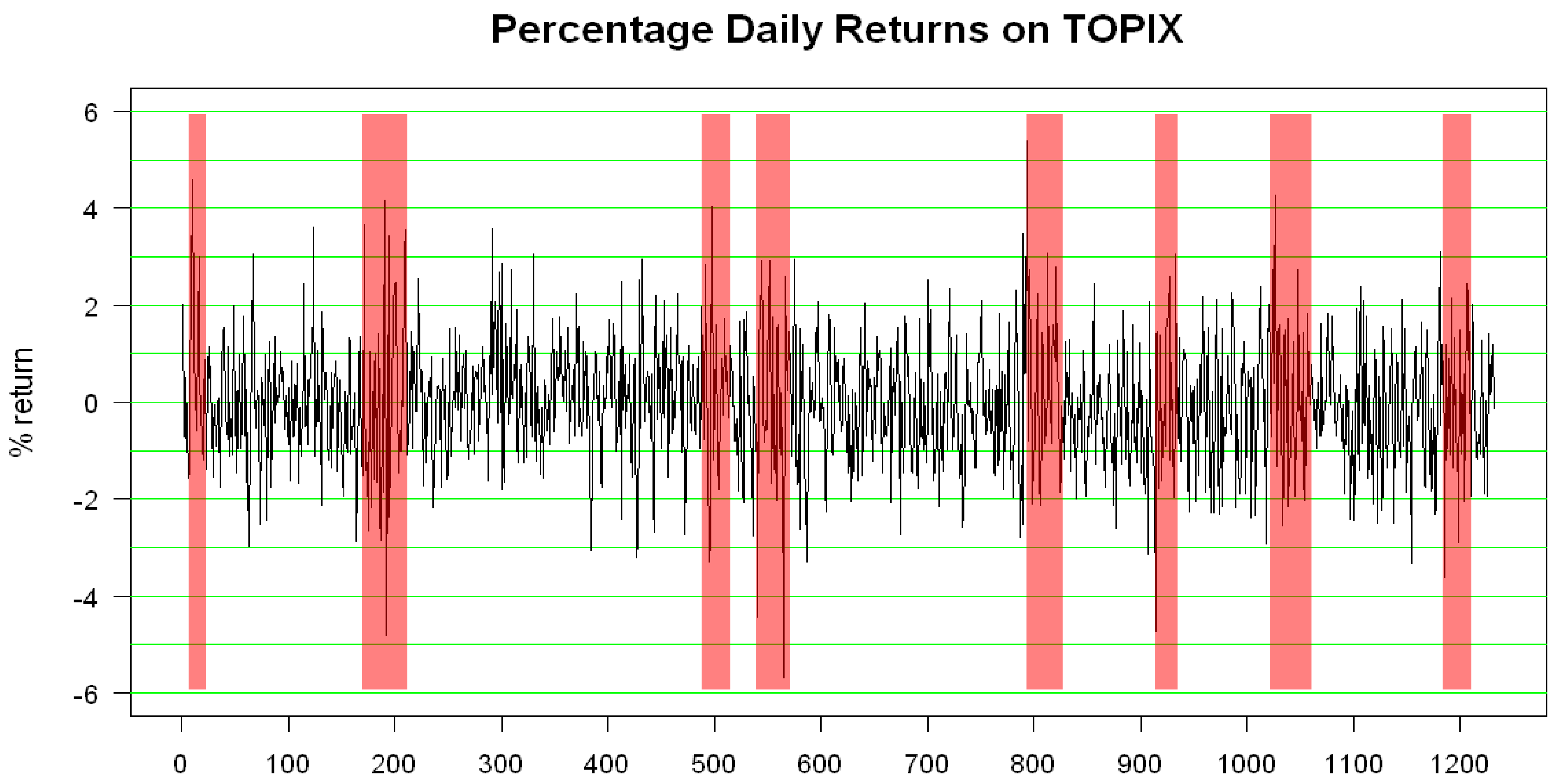
i	1	2	3	4
$\widehat{\delta}_i$	0.397	0.321	0.195	0.087
$\widehat{\mu}_i$	-0.0132	0.0037	-0.1273	-0.1179
$\widehat{\zeta}_{i0}$	0.917	0.000	0.000	0.564
$\widehat{\zeta}_{i1}$	31.01×10^{-5}	4.48×10^{-5}	9.61×10^{-5}	0.53×10^{-5}

Model 3(b) – state-dependent distributions for four-state von-Mises-HMM



Data:	Daily opening prices of the Tokyo Stock Price Index (TOPIX)
Period:	30.12.1997 – 30.12.2002
Length:	1233 trading days

Aim: Fit a Stochastic Volatility (SV) model to the daily returns



Periods of high volatility

The returns $y_t,\ t = 1, 2, \ldots, T$, on an asset satisfy:

$$y_t = \epsilon_t \beta e^{g_t/2} \qquad \epsilon_t \overset{\text{iid}}{\sim} N(0,1) \qquad \text{the observation process}$$

$$g_{t+1} = \phi g_t + \eta_t \qquad \eta_t \overset{\text{iid}}{\sim} N(0,\sigma^2) \qquad \text{an AR(1) state process}$$

SV without leverage: ϵ_t and η_t independent

parameters: $\phi,\ \sigma,\ \beta$

SV with leverage: $\text{cor}(\epsilon_t, \eta_t) = \rho$

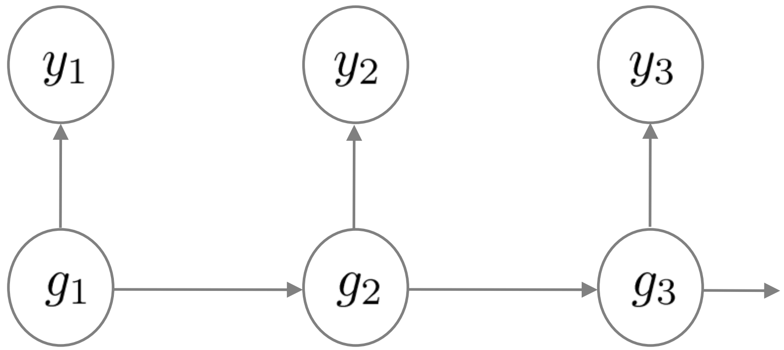
parameters: $\phi,\ \sigma,\ \beta\ ,\ \rho$

SV models are bedevilled by the difficulty of evaluating the likelihood.

They have an HMM structure.

observed

hidden



conditional independence
$y_t \sim N(0, \beta^2 e^{g_t})$

Markov property
$g_{t+1} = \phi g_t + \eta_t$

The state variable, g_t, is continuous, not discrete.
The likelihood is a T-fold integral which does not simplify.

Shephard (1996): "[There is a] vast literature on fitting SV models".

Trick: Discretize the state space into m states.

This results in a HMM with three parameters (β, ϕ, σ^2),
whose likelihood is easy to compute and maximize.

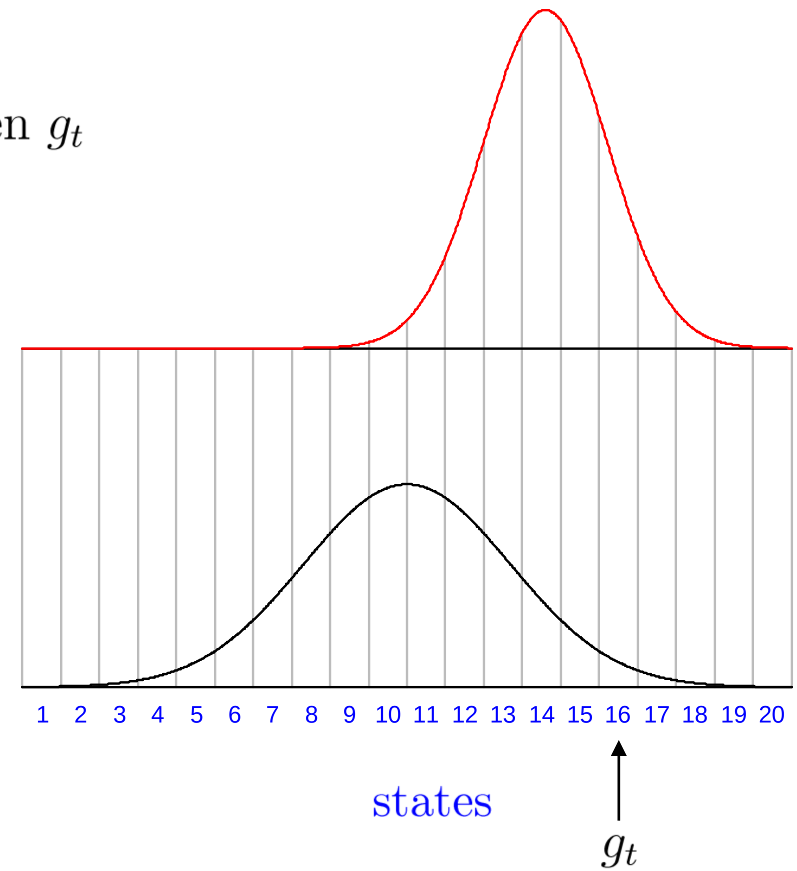
Approximation with $m = 20$ states

Conditional distribution of g_{t+1} given g_t

$g_{t+1} \sim N(\phi g_t, \sigma^2)$

Marginal distribution of g_t

$g_{t+1} \sim N(0, \frac{\sigma^2}{1-\phi^2})$



Transition probability matrix: $\Gamma : \gamma_{ij} = \Pr(g_{t+1} \in \text{state } j \mid g_t \in \text{ state } i)$

State-dependent distribution: $p(y_t|g_t)$ is $N(0, \beta^2 e^{g_t})$

Both are available in terms of ϕ, σ^2, and β.

A drop in return, y_t, is often followed by an increase in volatility g_{t+1}. (Cappé *et al.*, 2005, p. 28)

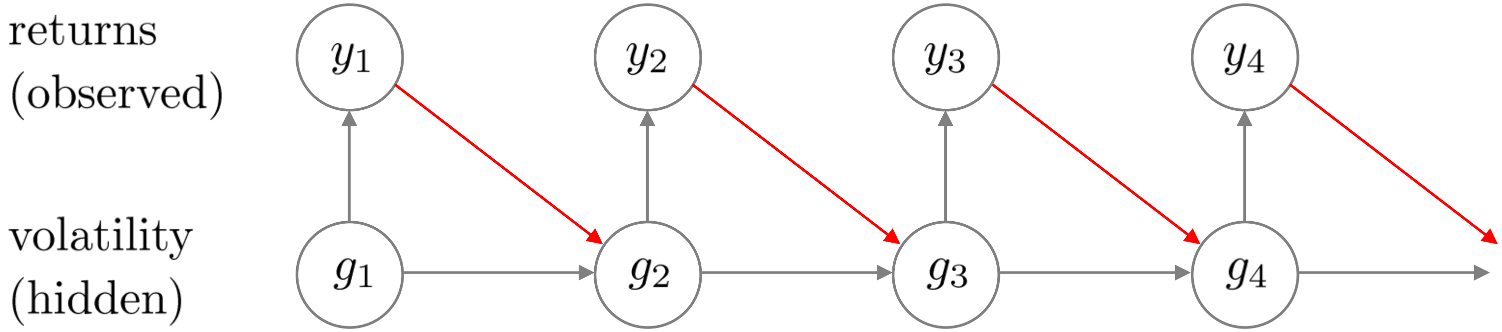
SV model with leverage

$$y_t = \epsilon_t \beta e^{g_t/2} \qquad \epsilon_t \overset{\text{iid}}{\sim} N(0,1)$$

$$g_{t+1} = \phi g_t + \eta_t \qquad \eta_t \overset{\text{iid}}{\sim} N(0,\sigma^2)$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix} \right)$$

There is **feedback** from past returns to volatility.



Application: TOPIX daily opening prices, 30.12.1997 – 30.12.2002

Omori, Y., Chib, S., Shephard, N. and Nakajima, J. (2007).
Stochastic volatility with leverage: fast and efficient likelihood inference.
Journal of Econometrics **140**, 425-449.

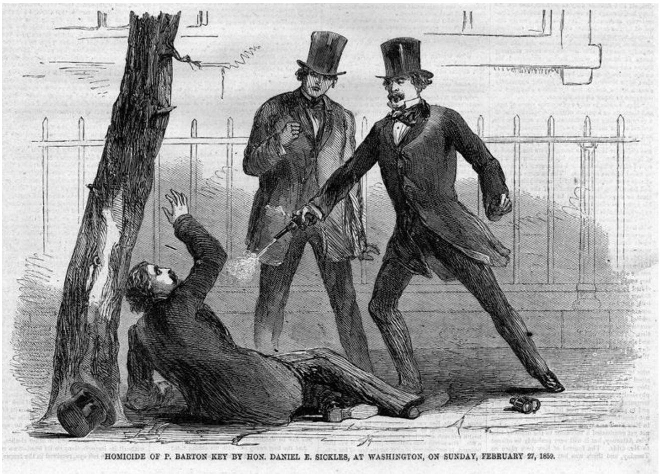
Estimates: HMM with m states and those[1] given in Omori *et al.* (blue)

m	ϕ	σ	β	ρ
10	0.935	0.129	1.206	-0.551
25	0.949	0.135	1.205	-0.399
50	0.949	0.140	1.205	-0.383
100	0.949	0.142	1.205	-0.379
200	0.949	0.142	1.205	-0.378
posterior means	0.951	0.134	1.205	-0.362

Confidence intervals: HMM case based on $m = 50$ and parametric bootstrap.

	ϕ	σ	β	ρ
95 % interval	(0.827, 0.973)	(0.078, 0.262)	(1.099, 1.293)	(-0.675, -0.050)
95 % interval	(0.908, 0.980)	(0.091, 0.193)	(1.089, 1.318)	(-0.593, -0.107)

[1]The unweighted version.



Source[1]:	S.A. Police mortuary, Salt River, Cape Town
Period:	01.01.1986 – 31.12.1991
Length:	5 series of length 313 weeks
Recorded:	Counts categorized as follows:

- firearm homicide (FH)
- nonfirearm homicide (NFH)
- firearm suicide (FS)
- nonfirearm suicide (NFS)
- legal intervention homicide (LIH)

Aims: Look for trends and patterns in these series, especially those relating to the use of firearms.
Was there a change in pattern at about 2 February 1990[2]?

[1] Source: Dr. L. B. Lerer; cf. MacDonald and Lerer (1994).

[2] Date of the speech by President F.W de Klerk that (effectively) ended apartheid.

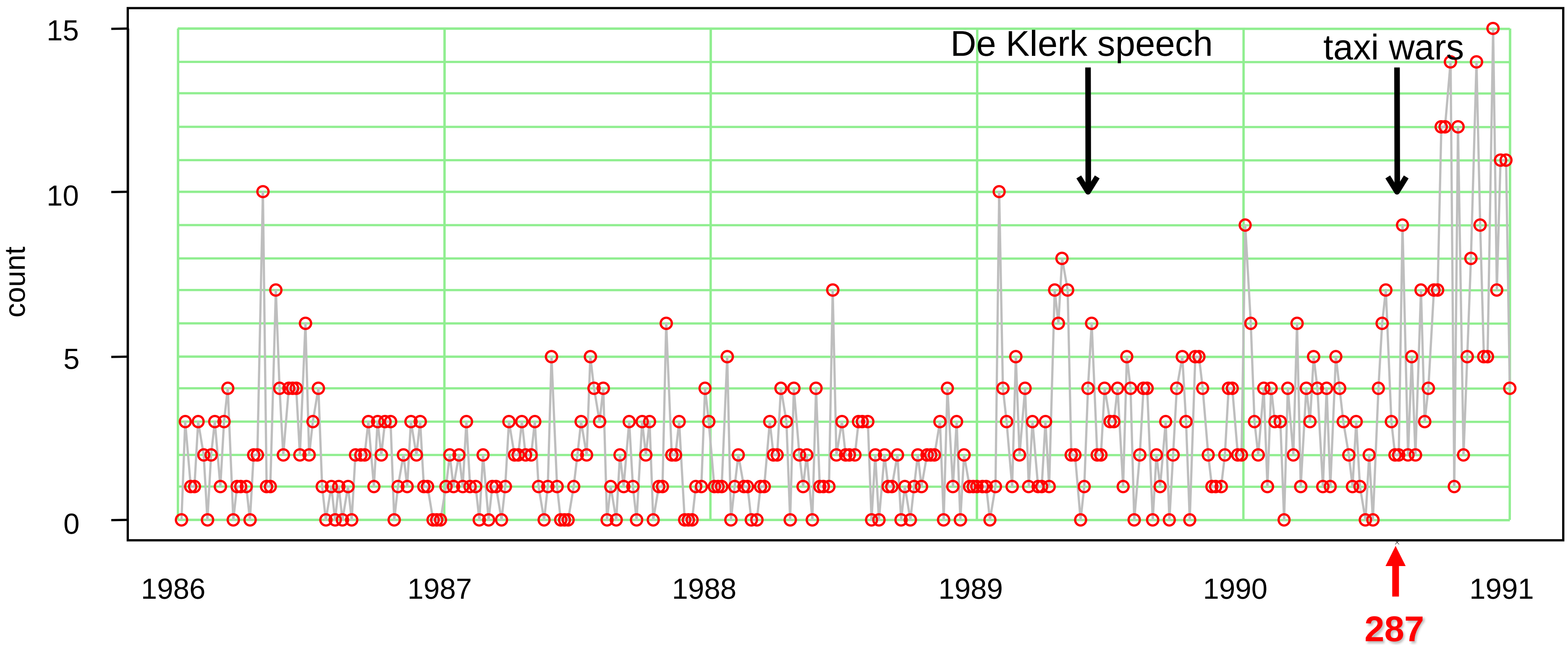
Time series examined

1. Counts — firearm homicides

2. Proportions: $\dfrac{\text{firearm homicides}}{\text{all homicides and suicides}}$

3. Multivariate model: proportions in each of the 5 categories

- firearm homicide	(FH)
- nonfirearm homicide	(NFH)
- firearm suicide	(FS)
- nonfirearm suicide	(NFS)
- legal intervention homicide	(LIH)



There was a marked increase firearm homicides in 1991.

The main cause was probably rapid urbanization at the time

⇒ dramatic increase in population in and around Cape Town

⇒ increase in population exposed.

Model	Poisson distribution parameters
• no trend	λ_1 and λ_2
• one time-trend parameter	$\log \lambda_1 = a_1 + b\,t,$ $\log \lambda_2 = a_2 + b\,t.$
• two time-trend parameters	$\log \lambda_1 = a_1 + b_1 t + b_2 t^2,$ $\log \lambda_2 = a_2 + b_1 t + b_2 t^2$
• change–point at time 287	$\lambda_i = \begin{cases} \lambda_i^{(1)} & \text{for } t < 287 \\ \lambda_i^{(2)} & \text{for } t \geq 287 \end{cases}$

Comparison of four HMM models

model with	k	L_T	AIC	BIC
λ_1 and λ_2 constant	4	626.64	1261.27	1276.26
loglinear trend	5	606.82	1223.65	1242.38
log-quadratic trend	6	602.27	1216.55	1239.02
change-point at time 287	6	605.56	1223.12	1245.60

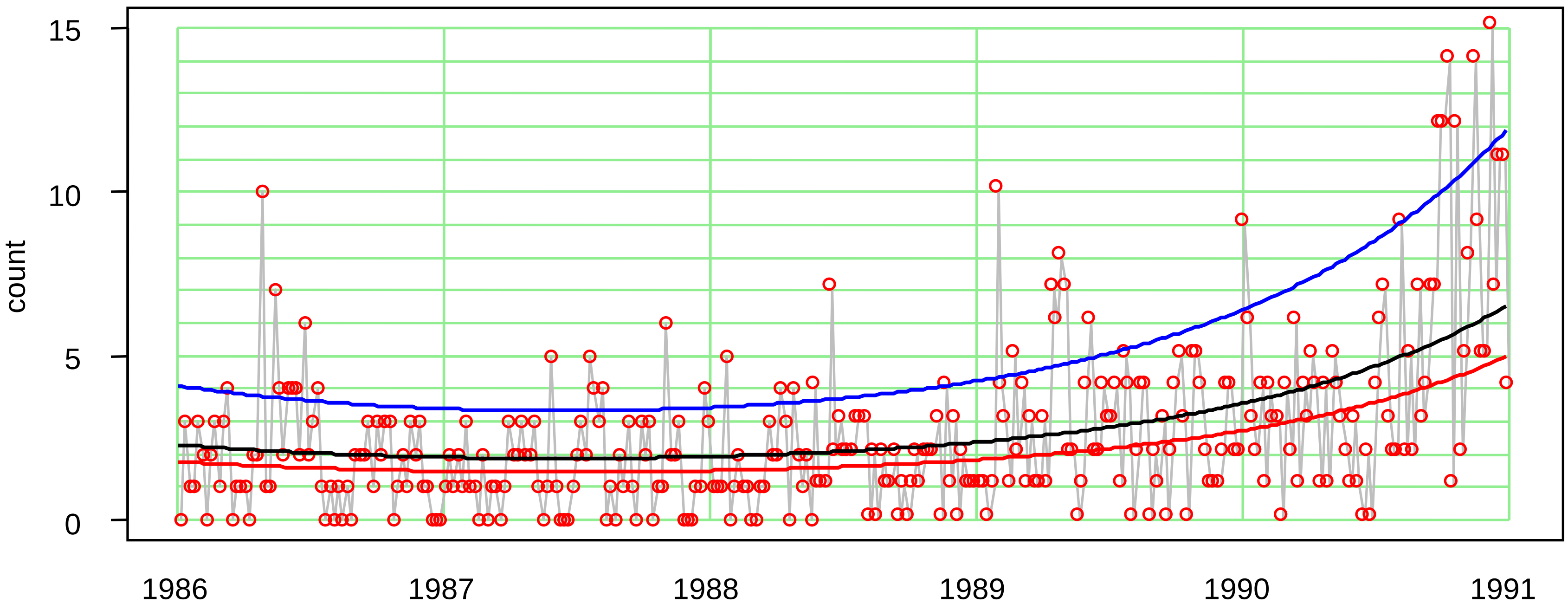
Two-state Poisson-HMM with quadratic trends

$$\hat{\delta}' = \begin{pmatrix} 0.777 \\ 0.223 \end{pmatrix} \qquad \hat{\Gamma} = \begin{pmatrix} 0.881 & 0.119 \\ 0.416 & 0.584 \end{pmatrix}$$

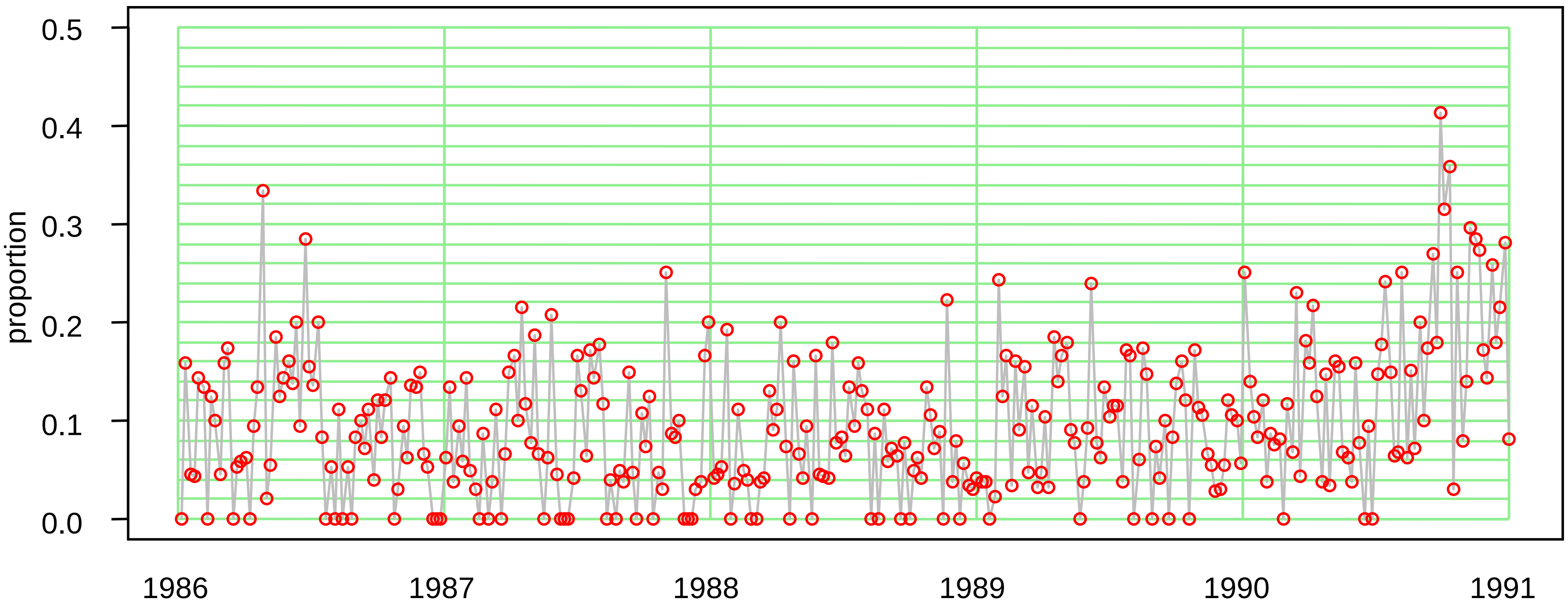
State 1 relatively persistent

$$\widehat{\log \lambda_1} = 0.477 - 0.004858\, t + 0.00002665\, t^2$$
$$\widehat{\log \lambda_2} = 1.370 - 0.004858\, t + 0.00002665\, t^2$$

Weekly counts of firearm homicides and fitted state-dependent trends



Firearm homicides as a proportion of all homicides and suicides



There was a marked increase in the **proportion** of firearm homicides in 1991. This cannot be attributed purely to an increase in the population; other causes need to be considered.

Two-state binomial-HMMs

- Let n_t be the total number of homicides and suicides in week t,

Model	Parameters of binomial distribution
• no trend	π_1 and π_2
• one time-trend parameter	$logit\ \pi_1 = a_1 + b\,t,$ $logit\ \pi_2 = a_2 + b\,t.$
• two time-trend parameters	$logit\ \pi_1 = a_1 + b_1 t + b_2 t^2,$ $logit\ \pi_2 = a_2 + b_1 t + b_2 t^2$
• change-point at time 287	$\pi_i = \begin{cases} \pi_i^{(1)} & \text{for } t < 287 \\ \pi_i^{(2)} & \text{for } t \geq 287 \end{cases}$

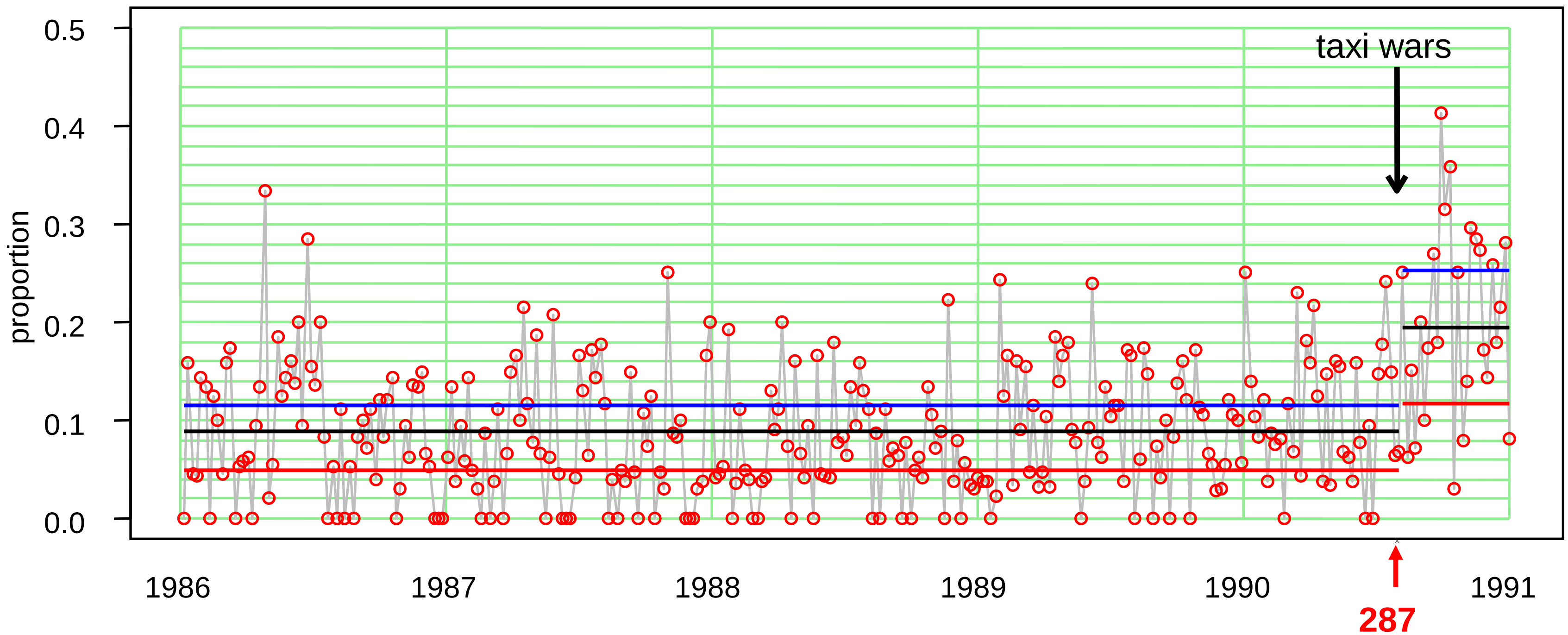
Model selection criteria

model with	k	$-\mathrm{L}_T$	AIC	BIC
π_1 and π_2 constant	4	590.26	1188.52	1203.50
one time-trend parameter	5	584.34	1178.67	1197.40
two time-trend parameters	6	581.87	1175.75	1198.23
change-point at time 287	6	573.27	1158.55	1181.03

$$\hat{\delta}' = \begin{pmatrix} 0.426 \\ 0.574 \end{pmatrix} \qquad \hat{\Gamma} = \begin{pmatrix} 0.658 & 0.342 \\ 0.254 & 0.746 \end{pmatrix}$$

Neither state is very persistent.

$$\pi_1^{(1)} = 0.050 \quad \pi_2^{(1)} = 0.116 \quad \text{for } t < 287$$
$$\pi_1^{(2)} = 0.117 \quad \pi_2^{(2)} = 0.253 \quad \text{for } t \geq 287$$



Categories:

- firearm homicide (FH)
- nonfirearm homicide (NFH)
- firearm suicide (FS)
- nonfirearm suicide (NFS)
- legal intervention homicide (LIH)

Observations: $x_t = (x_{t1}, x_{t2}, x_{t3}, x_{t4}, x_{t5}), \quad t = 1, 2, \ldots, T,$

where $n_t = \sum_{j=1}^{5} x_{tj}$ is regarded as *given*.

Likelihood: $\delta\, \mathrm{P}(x_1)\Gamma\, \mathrm{P}(x_2)\Gamma\, \mathrm{P}(x_3) \cdots \Gamma\, \mathrm{P}(x_T)1'$

where $\mathrm{P}(x_t)$ is a diagonal matrix with i–th entry $\Pr(X_t = x_t |\, C_t = i)$.

$$\Pr(X_t = x_t |\, C_t = i) \;=\; \binom{n_t}{x_{t1}\ x_{t2}\ x_{t3}\ x_{t4}\ x_{t5}}\, \pi_{i1}^{x_{t1}} \pi_{i2}^{x_{t2}} \pi_{i3}^{x_{t3}} \pi_{i4}^{x_{t4}} \pi_{i5}^{x_{t5}}$$

To model a change-point

Simply use one set of parameters for the state-dependent model before the change-point, and another set after the change-point.

$$\hat{\Gamma} = \begin{pmatrix} 0.541 & 0.459 \\ 0.097 & 0.903 \end{pmatrix} \qquad \hat{\delta}' = \begin{pmatrix} 0.174 \\ 0.826 \end{pmatrix}$$

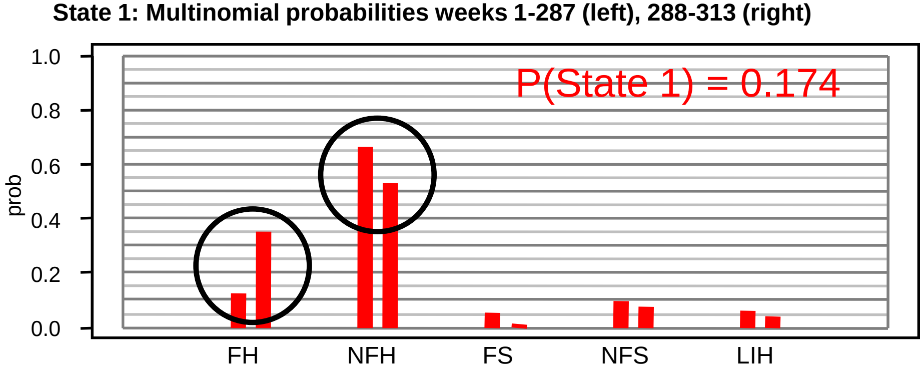
State–dependent distributions

Weeks 1–287

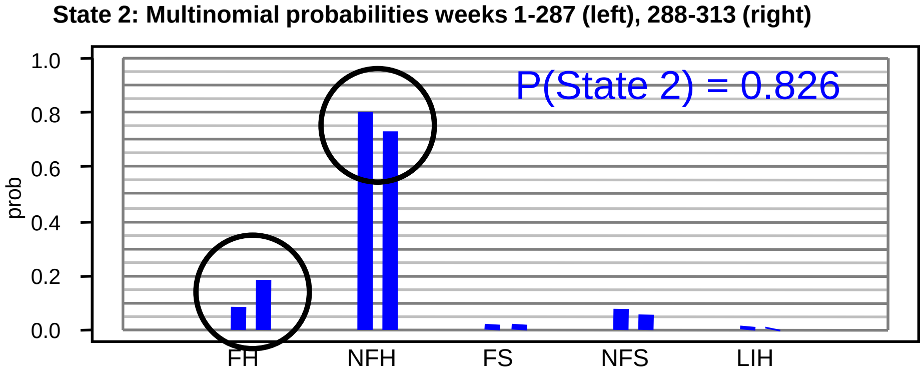
	FH	NFH	FS	NFS	LIH
in state 1	0.124	0.665	0.053	0.098	0.059
in state 2	0.081	0.805	0.024	0.074	0.016
marginal	0.089	0.780	0.029	0.079	0.023

Weeks 288–313

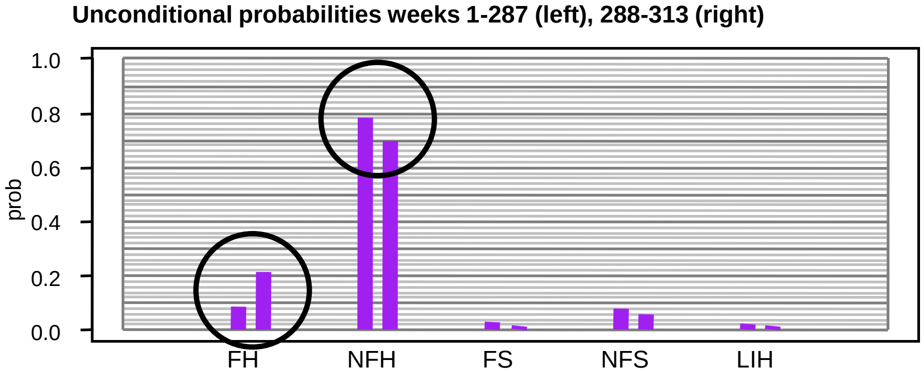
	FH	NFH	FS	NFS	LIH
in state 1	0.352	0.528	0.010	0.075	0.036
in state 2	0.186	0.733	0.019	0.054	0.008
marginal	0.215	0.697	0.018	0.058	0.013



There was a shift from NFH to FH. However, the process is relatively infrequently in state 1.



There was a shift from NFH to FH.



The shift is (of course) evident in the unconditional distribution.

Hidden Markov models are

- **satisfyingly flexible,**
- **reasonably easy to apply,**
- **moderately parsimonious,**
  A 3–state model often provides a reasonable fit.
- **occasionally interpretable.**

Azzalini, A. and Bowman, A.W. (1990). A look at some data on the Old Faithful geyser. *Appl. Statist.* **39**, 357–365.

Cappé, O., Moulines, E. and Rydén, T. (2005). *Inference in Hidden Markov Models.* Springer, New York.

MacDonald, I.L. and Lerer, L.B. (1994). A time series analysis of trends in firearm-related homicide and suicide. *Int. J. Epidemiol.* **23**, 66–72.

MacDonald, I.L. and Zucchini, W. (1997). *Hidden Markov and Other Models for Discrete-valued Time Series.* Chapman & Hall, London.

Omori, Y., Chib, S., Shephard, N. and Nakajima, J. (2007). Stochastic volatility with leverage: fast and efficient likelihood inference. *J. Econometrics* **140**, 425–449.

Shephard, N.G. (1996). Statistical Aspects of ARCH and stochastic volatility. In *Time Series Models: In econometrics, finance and other fields*, D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen (eds), 1–67. Chapman & Hall, London.