UDC 519.246.8:616.379-008.64

MARKOV-SWITCHING VECTOR AUTOREGRESSION FOR LATENT GLYCAEMIC REGIMES IN REAL-WORLD DIABETES SELF-MONITORING DATA: AN OUT-OF-SAMPLE EVALUATION

[1]*Isomiddinov A.I.,* [1*]*Akbarova A.A.,* [2]*Akbarov B.X.*

[*]akbarovaaa1991@gmail.com

[1]Namangan State Technical University, 12, Kosonsoy str., Namangan, 160115 Uzbekistan;
[2]Kimyo International University in Tashkent, 156, Usman Nasyr str., Tashkent, 100121 Uzbekistan

Self-monitoring records of blood glucose and insulin dose form irregular time series whose physiological regime is not observed. This study fits a Markov-switching vector autoregression, MS-VAR(1), with regime-specific intercepts, autoregressive matrices and covariances to raw event logs of the public UCI AIM'94 data and asks what regime switching adds to out-of-sample prediction. Each patient is modelled separately by expectation–maximisation with exact forward–backward smoothing, without any regime label; 20 of 70 patients (2189 glucose-recording days) met a priori recording criteria. Models were compared by one-step-ahead predictive log-likelihood per day on the final 20% of each series against a single-regime VAR(1), a Gaussian hidden Markov model (HMM), a single-pass switching model and a static Gaussian, using paired Wilcoxon tests with Holm correction and BIC. A feature set containing both levels and first differences is degenerate for a first-order autoregression: two features are exact linear functions of the lagged state, and the smallest residual-covariance eigenvalue equalled the regularisation floor in all 20 patients. In that specification the advantage over the HMM was reproduced, but a single-regime VAR(1) did equally well. After removing the redundant features, MS-VAR(3) was not better than the single-regime VAR(1) (median paired difference -0.31 nats/day, 95% CI $[-1.11, 0.89]$), BIC selected at most two regimes in 17 of 20 patients, and the advantage over the HMM ($+2.33$ nats/day) did not survive Holm correction ($p = 0.069$). The fitted regimes were persistent (mean dwell time 10.2 days) and ordered in mean glucose (137, 160, 194 mg/dL), but day-level labels were noise-sensitive (ARI 0.78 at noise SD 0.05). Regime switching is thus supported as a descriptive segmentation, not as a predictive improvement.

Keywords: diabetes self-monitoring data, Markov-switching vector autoregression, hidden Markov model, latent regime, out-of-sample evaluation, held-out log-likelihood, model selection

Citation: Isomiddinov A.I., Akbarova A.A., Akbarov B.X. 2026. Markov-switching vector autoregression for latent glycaemic regimes in real-world diabetes self-monitoring data: an out-of-sample evaluation. *Problems of Computational and Applied Mathematics.* 5(76): 60-71.

DOI: 10.71310/pcam.5_76.2026.05

1 Introduction

Diabetes is managed day to day through self-monitoring records: blood-glucose measurements and insulin doses collected as irregular time series in routine outpatient care. These records are unevenly sampled, mix electronic and paper logging, and carry no label

of the patient's physiological regime. Recovering a regime from them is a latent-variable problem, not a supervised classification task.

Three levels of temporal modelling are common. Static clustering groups similar days but treats each day independently. A hidden Markov model (HMM) [1] adds persistence of a discrete regime but treats, given the regime, each day as unrelated to the previous one. A Markov-switching autoregression [2, 3] also lets the day-to-day dynamics depend on the regime, and a switching state-space model [4] adds a continuous latent state. Whether the added structure pays off on real self-monitoring data is empirical and cannot be settled by classification accuracy, since no ground-truth regime exists. The question of this paper is therefore: *what does regime switching add to out-of-sample prediction of daily glucose–insulin records, beyond persistence alone and beyond a single-regime autoregression?*

Related work. Physiological modelling and control in diabetes are reviewed in [5]; the minimal model [6] describes continuous glucose–insulin states, and staged β-cell dysfunction [7] motivates discrete regimes only in general terms, since neither insulin sensitivity nor β-cell mass is identifiable from the data used here. Kalman-type filters [8] have been applied to glucose dynamics from continuous monitoring [9] or simulation [10], a setting different from sparse daily records. Switching models are estimated by EM [11], with filtering and smoothing recursions in [12]. Evaluations of such models are mostly analytical or simulated with known regimes; here the raw AIM'94 event logs [13] are used and evaluation rests on held-out likelihood, information criteria and stability analyses.

Contribution. The paper proposes no new estimator but a designed evaluation: (i) a reproducible pipeline from raw event logs to daily features with every exclusion documented; (ii) baselines that isolate persistence, autoregression, regime-specific dynamics and iterative refinement; (iii) diagnosis of a degenerate feature specification (levels together with first differences) that creates a spurious autoregressive advantage, with the corrected analysis; (iv) a descriptive, non-clinical characterisation of the regimes.

2 Data and cohort

The AIM'94 data [13] are 70 tab-separated files (1988–1991), each row holding a date, a time, an event code and a value, recorded by an electronic device or on paper (with fixed logical times). Codes 33–35 are regular, NPH and UltraLente insulin; codes 48 and 57–64 are glucose measurements; codes 65–72 are non-numeric events not used. Lines without exactly four fields or a valid date were excluded and counted. Exclusions reported explicitly: seven lines dated 06-31-1991 in data-20 (non-existent date); 3 and 30 corrupted line pairs in data-27 and data-29; seven out-of-range meter readings ("0Hi", "0Lo") in data-02, dropped from daily means and not replaced by numbers; and code 56, absent from the code table, present in 24 files (11 times in data-50), excluded from both series. No non-numeric value was treated as zero, and glucose values were not screened for plausibility beyond numeric parsing.

Recording span is the inclusive number of calendar days between a patient's first and last valid record of any kind; a *glucose-recording day* has at least one valid glucose reading. Criteria fixed a priori were span $\geqslant 60$ days and $\geqslant 40$ glucose-recording days. They retain 20 patients with 2189 glucose-recording days (58–166 per patient, median 119; Table 1); patient 70 qualifies because its last three days contain only insulin entries. Training segments contain 44–131 one-step transitions per patient and held-out segments 12–33.

Glucose readings on a date are averaged to a daily mean G_t (daily minimum and maximum are kept for the descriptive analysis) and doses are summed to I_t; a day without logged dose has $I_t = 0$, meaning no logged dose, not no insulin. The index t counts glucose-

Table 1 Cohort: 20 of 70 AIM'94 patients (original file numbers).

Patient	Glucose days	Span	Patient	Glucose days	Span
68	166	288	29	159	232
55	149	149	30	148	152
65	137	137	01	136	136
20	134	135	56	125	126
28	122	171	27	119	272
67	119	120	54	103	103
02	96	96	31	85	88
26	79	199	42	66	66
21	64	65	41	63	63
50	61	61	70	58	61

recording days, not calendar days; no missing day is imputed. The main feature vector is

$$y_t = \left[G_t,\ I_t,\ \bar{G}_t,\ \bar{I}_t\right]^{\mathrm{T}}, \tag{1}$$

where $\bar{G}_t, \bar{I}_t$ are trailing means over the $\min(t, 7)$ latest recorded days (strictly causal). Each component is standardised per patient (training segment only in held-out evaluation); constant features are dropped (patient 50 has no logged insulin). Two other specifications are analysed: a *six-feature* one adding $\Delta G_t = G_t - G_{t-1}$ and $\Delta I_t = I_t - I_{t-1}$, and a *two-feature* one (G_t, I_t).

3 Model and estimation

3.1 From switching state-space model to MS-VAR

A general switching state-space model couples a regime $s_t \in \{1, \ldots, M\}$ with a continuous state x_t: $x_{t+1} = f_{s_t}(x_t) + w_t$, $y_t = h_{s_t}(x_t) + v_t$ [4]. Fitting it needs known regimes or an informative prior on h; AIM'94 has no independent measurement of β-cell function or insulin sensitivity, so $x_t \neq y_t$ is unidentifiable. With the identity operator the continuous state disappears and a Markov-switching vector autoregression (MS-VAR) remains [2, 3]; the model is therefore called MS-VAR(1) and is not claimed to be a state-space model with a physiological latent state:

$$y_t = c_{s_t} + A_{s_t} y_{t-1} + w_t, \quad w_t \sim \mathcal{N}(0, Q_{s_t}), \quad P(s_t = j \mid s_{t-1} = i) = p_{ij}. \tag{2}$$

The intercept c_j gives each regime its own mean level; without it every regime would share the patient-level mean, contradicting the aim of identifying glycaemic regimes. The number of free parameters is $\varkappa = M[d^2 + d + d(d+1)/2] + M(M-1) + (M-1)$, i.e. 98 for $d = 4$, $M = 3$, against 44–131 training transitions per patient (median 93).

3.2 Why levels and first differences cannot be combined

If y_t contains G_t and ΔG_t, then $G_t - \Delta G_t = G_{t-1}$ is exactly the first coordinate of y_{t-1}, and likewise for insulin. In a VAR(1) this combination is predicted without error in every regime, so the unregularised residual covariance is singular in two directions and the likelihood is unbounded. With covariance floor ε, each such direction adds up to $-\frac{1}{2}\ln(2\pi\varepsilon)$ nats per step (1.38 at $\varepsilon = 10^{-2}$; 2.77 for both), an amount set by the regularisation constant, not by the data. An i.i.d.-emission HMM cannot exploit this identity and is penalised relative to any autoregressive model, whatever the regime structure. The main specification therefore uses levels and moving averages only. The moving averages are not

exactly redundant, but they are autocorrelated by construction (consecutive values share six of seven observations); this matters for interpretation (Section 3).

3.3 Expectation–maximisation

Each patient's model is fitted by EM [11] (Algorithm 1). The first observed step uses the initial distribution ν in place of a predicted distribution. Initialisation uses K-means ($k = M$, ten starts [14]); K-means plays no further role. A regime with total smoothed weight below $3d$ ($2d$ for the HMM) keeps its previous parameters, and Gaussian log-densities use Cholesky factorisation with jitter only on failure. Because the ridge on A_j and the floor on Q_j do not correspond to the exact maximiser of a penalised objective, monotone increase of the likelihood is not guaranteed; it was checked rather than assumed: 70% of MS-VAR(3) fits increased monotonically, with a largest one-iteration decrease of 0.034 nats. The procedure is regularised maximum likelihood and is not Bayesian. Regime labels are exchangeable: within-fit metrics do not depend on labelling, regimes are ranked post hoc by mean glucose within each patient for description, and the adjusted Rand index (ARI) [15] is permutation-invariant.

Algorithm 1 MS-VAR(1) estimation by EM

Require: standardised sequence $\{y_t\}_{t=2}^{T}$, M, $\lambda = \varepsilon = 10^{-2}$, tolerance 10^{-5}, at most 100 iterations.

Ensure: $\Theta^* = \{P, \nu, \{c_j, A_j, Q_j\}\}$, smoothed $\gamma_t(j)$, labels $\hat{s}_t = \arg\max_j \gamma_t(j)$.

1: Initialise labels by K-means; obtain ν, P (rows floored at 10^{-6}), and c_j, A_j, Q_j by hard-weighted least squares.

2: **repeat**

3: **E-step:** scaled forward–backward in log space with $L_t(j) = \mathcal{N}(y_t;\, c_j + A_j y_{t-1}, Q_j)$; obtain filtered $\alpha_t(j)$, smoothed $\gamma_t(j) = P(s_t = j \mid y_{2:T})$, $\xi_t(i,j) = P(s_{t-1} = i, s_t = j \mid y_{2:T})$.

4: **M-step:** $p_{ij} = \sum_t \xi_t(i,j) \big/ \sum_t \sum_{j'} \xi_t(i,j')$; with $z_t = [y_{t-1}; 1]$,

$$[A_j\ c_j] = \Big(\sum_t \gamma_t(j)\, y_t z_t^{\mathrm{T}}\Big)\Big(\sum_t \gamma_t(j)\, z_t z_t^{\mathrm{T}} + \mathrm{diag}(\lambda \mathbf{I}, 0)\Big)^{-1}, \tag{3}$$

$$Q_j = \frac{\sum_t \gamma_t(j)\, r_t r_t^{\mathrm{T}}}{\sum_t \gamma_t(j)} + \varepsilon \mathbf{I}, \qquad r_t = y_t - c_j - A_j y_{t-1}. \tag{4}$$

5: **until** relative change of the log-likelihood $< 10^{-5}$ or 100 iterations

4 Evaluation design

Models (per patient, same observations, $M = 3$ unless stated; each baseline isolates one component): static Gaussian ($M = 1$, i.i.d.); Gaussian HMM (persistence without autoregression); *single-regime VAR(1)* (autoregression without regimes, the key baseline); single-pass switching model (K-means partition taken as final, VAR(1) fitted once with hard weights, isolating iterative refinement); and MS-VAR(3). For internal validity, K-means, PCA (two components) + K-means and a shallow autoencoder (tanh, hidden layers 3–2–3) + K-means are added; the single-pass model shares K-means' partition and indices.

Held-out likelihood. Each series is split in temporal order at 80% (at least 10 training observations); standardisation, K-means and EM use the training part only. On the held-out part the filter continues from the final filtered state, the first prediction uses the last training observation as y_{t-1}, and the score is $\sum_t \log \sum_j \alpha_{t|t-1}(j) L_t(j)$ *divided by the number of held-out steps*, since segments differ in length and sums are not comparable between patients. HMM and static models use the density of y_t without conditioning on y_{t-1}.

Statistics. The unit of analysis is the patient ($N = 20$). Paired comparisons use the two-sided Wilcoxon signed-rank test, the matched-pairs rank-biserial r, the median difference and a bootstrap 95% CI (5000 resamples). Holm's correction [16] is applied within declared families: the four comparisons of MS-VAR(3) with the HMM, single-pass, VAR(1) and static baselines (primary family, and likewise within each sensitivity configuration), the three decomposition contrasts, and the four silhouette comparisons; raw and adjusted p are reported. The number of regimes is compared by held-out likelihood and by training-segment BIC [17]. Stability is assessed over ten seeds (0, 1, 2, 3, 4, 7, 21, 42, 50, 100) and over additive Gaussian noise (SD 0–0.30, five draws), with ARI against the unperturbed fit as a within-model reference, not a ground truth. Internal validity uses the silhouette [18], Calinski–Harabasz [19] and Davies–Bouldin [20] indices. Regimes are compared descriptively with two operational thresholds (daily minimum < 70 and maximum > 180 mg/dL).

5 Results

5.1 Diagnosis of the six-feature specification

For the single-regime VAR(1) fitted to training segments, the smallest eigenvalue of the residual covariance with six features was within 50% of the regularisation floor in 20 of 20 patients (median 0.0100) and the second smallest in 19 of 20; with four features the counts were 2 and 0 (medians 0.035 and 0.103). The covariance of the six-feature model is thus controlled by the regularisation constant, as predicted in Section 2. In this specification autoregressive models reach a far higher held-out likelihood than the HMM and static model (Table 2), including positive total values, and MS-VAR(3) beats the HMM in 19 of 20 patients (median +7.12 nats/day, Holm $p = 0.003$). But MS-VAR(3) is indistinguishable from the single-regime VAR(1) (+0.06, Holm $p = 0.46$) and the single-pass model (Table 3). The advantage over the HMM is therefore a consequence of autoregression on features with an exact lagged identity, not of regime switching, and this specification is not used further for conclusions.

Table 2 Held-out log-likelihood per day (median [IQR] over 20 patients; mean ± SD for the main specification only). Six-feature: levels, moving averages, first differences, no intercept; main: $G, I, \bar{G}, \bar{I}$ with intercepts.

Model	Six-feature	Main	Main, mean ± SD	Params
Static Gaussian (M=1)	−8.38 [−12.25, −6.54]	−6.29 [−8.74, −4.69]	-14.5 ± 26.0	14
Gaussian HMM (M=3)	−9.30 [−13.23, −6.40]	−6.70 [−9.39, −4.56]	-15.6 ± 29.1	50
Single-regime VAR(1)	−1.21 [−3.28, −0.41]	−3.31 [−4.81, −2.56]	-18.6 ± 59.8	30
Single-pass switching (M=3)	−0.80 [−3.32, 0.10]	−4.09 [−5.51, −2.59]	-20.1 ± 60.3	98
MS-VAR(3), EM	−1.23 [−2.62, 0.21]	−3.30 [−6.02, −2.52]	-21.6 ± 61.8	98

5.2 Held-out prediction in the main specification

MS-VAR(3) attains a median held-out log-likelihood of −3.30 nats/day, equal to the single-regime VAR(1) (−3.31) and higher than the HMM (−6.70) and static Gaussian

(−6.29) (Table 2, Fig. 1). The paired tests (Table 3) show no difference from the single-regime VAR(1) (−0.31 nats/day, 95% CI [−1.11, 0.89]; MS-VAR better in 9 of 20 patients) or from the single-pass model, and the nominal advantages over the HMM (raw $p = 0.017$) and the static Gaussian (raw $p = 0.036$) are not significant after Holm correction (Holm $p = 0.069$ and 0.11). MS-VAR(3) has 98 free parameters; its mean training log-likelihood (−0.48 per step) far exceeds its held-out median (−3.30), which indicates over-fitting. The means in Table 2 are dominated by patients 21 and 31, whose held-out segments lie far outside the training range (for patient 31 the training SD of the insulin dose is 0.73 units and held-out doses are up to 56 training SDs below the mean); this affects all models. Individual values are shown in Fig. 1.

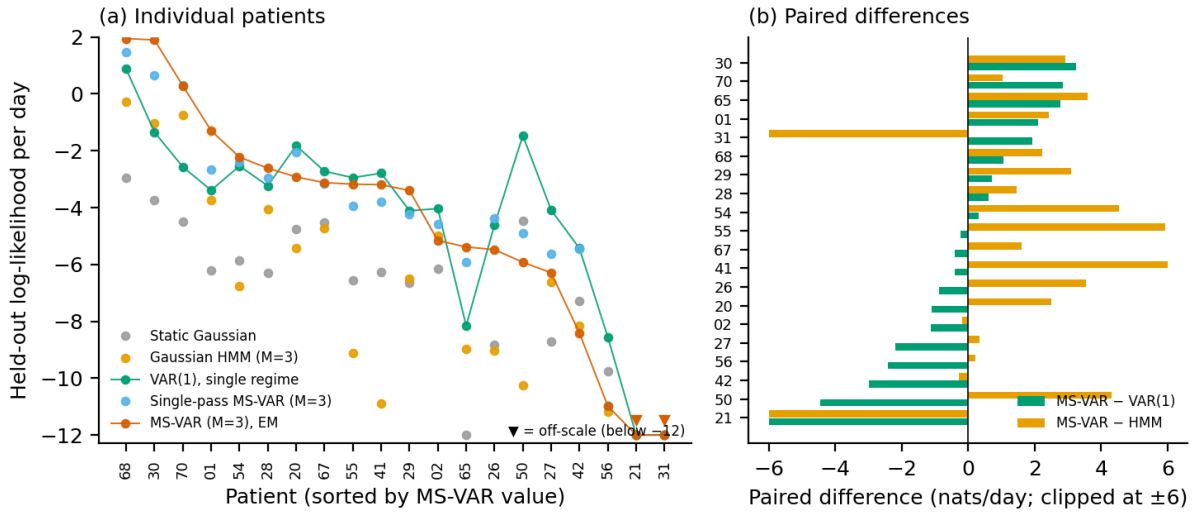


Figure 1 Held-out log-likelihood per day, main specification. (a) Individual patients (values below −12 drawn at the lower edge). (b) Per-patient paired differences of MS-VAR(3) from the single-regime VAR(1) and from the HMM (clipped at ±6).

Table 3 Median paired difference in held-out log-likelihood per day, MS-VAR(3) minus comparator (nats/day; positive favours MS-VAR). Main columns: 95% CI, Wilcoxon p, Holm p, patients where MS-VAR is better; other columns: difference (Holm p) in the six-feature, two-feature and $\lambda = \varepsilon = 10^{-3}$, 10^{-1} configurations.

	Main specification				Six-feat.	Two-feat.	10^{-3}	10^{-1}
Comparator	Diff. [95% CI]	p	Holm p	Wins				
vs Gaussian HMM	+2.33 [0.68, 3.33]	0.017	0.069	16/20	+7.12 (0.003)	−0.02 (1.00)	+2.11 (0.077)	+1.80 (0.003)
vs single-pass switching	+0.03 [−0.61, 0.51]	1.00	1.00	12/20	+0.08 (0.73)	+0.04 (1.00)	+0.01 (0.84)	+0.11 (0.40)
vs single-regime VAR(1)	−0.31 [−1.11, 0.89]	0.67	1.00	9/20	+0.06 (0.46)	+0.20 (1.00)	−0.82 (0.66)	+0.18 (0.84)
vs static Gaussian	+3.17 [1.20, 3.67]	0.036	0.11	15/20	+7.63 (0.003)	+0.04 (1.00)	+2.68 (0.19)	+2.95 (0.003)

5.3 Sensitivity and the source of the predictive gain

No configuration shows a significant difference between MS-VAR(3) and the single-regime VAR(1) or the single-pass model (Table 3). The advantage over the HMM and static baselines appears only when moving-average features are present and is significant after correction only for the strongest regularisation (+1.80 nats/day, Holm $p = 0.0028$); with the two raw features it is absent. A decomposition (exploratory, outside the primary family) explains the pattern. With the main features, persistence alone (HMM vs static)

gives no gain (-0.02, $p = 0.78$), whereas autoregression gives a large one (VAR(1) vs static $+2.97$, Holm $p = 0.0021$). With two raw features the autoregressive gain disappears ($+0.02$, $p = 0.31$) and only a weak, non-significant gain from persistence remains ($+0.77$, Holm $p = 0.21$). The autoregressive gain in the main specification therefore stems from the autocorrelated 7-day averages, not from day-to-day dynamics of glucose or insulin.

5.4 Number of regimes

Table 4 Number of regimes M ($M = 1$ is the single-regime VAR(1)): parameters and training log-likelihood per step (main specification), held-out log-likelihood per day (median [IQR]), and number of patients for whom M is best by BIC / by held-out likelihood.

			Main		Two features	
M	Params	Train LL	Held-out LL	BIC/held-out	Held-out LL	BIC/held-out
1	30	−2.32	−3.31 [−4.81, −2.56]	9/5	−2.99 [−4.79, −2.46]	6/5
2	63	−0.98	−3.02 [−6.84, −2.08]	8/3	−2.79 [−6.32, −1.98]	8/7
3	98	−0.48	−3.30 [−6.02, −2.52]	3/4	−3.10 [−4.16, −2.23]	5/2
4	135	−0.41	−2.99 [−5.61, −2.19]	0/8	−3.64 [−4.49, −2.18]	1/6

Training likelihood grows steadily with M, whereas held-out likelihood is flat (Table 4). BIC selects $M \leqslant 2$ in 17 of 20 patients and $M = 3$ in only 3; with two raw features the numbers are 14 and 5. The held-out criterion does not favour a consistent M. $M = 3$ is thus not supported over $M = 1$ or 2 by either criterion; it was kept, as fixed a priori, for the descriptive analyses below, which are descriptive segmentations under an assumed M.

5.5 Internal validity and temporal stability

Table 5 Internal validity and temporal stability, main specification, full-series fits (mean ± SD over 20 patients).

Method	Silhouette	Calinski–Harabasz	Davies–Bouldin	Dwell (days)	Transitions
K-means (= single-pass)	0.349 ± 0.071	63.3 ± 24.9	1.04 ± 0.22	4.7 ± 3.1	28.1 ± 14.1
PCA + K-means	0.430 ± 0.061	105.6 ± 43.7	0.78 ± 0.15	3.3 ± 0.9	32.8 ± 11.7
Autoencoder + K-means	0.535 ± 0.058	155.3 ± 51.4	0.65 ± 0.13	3.7 ± 1.3	30.8 ± 12.6
Gaussian HMM	0.309 ± 0.105	40.1 ± 18.3	1.35 ± 0.36	12.7 ± 8.0	10.9 ± 9.0
MS-VAR(3), EM	0.233 ± 0.153	33.9 ± 21.3	1.65 ± 0.92	10.2 ± 8.8	17.7 ± 16.3

MS-VAR(3) has the lowest silhouette and the highest Davies–Bouldin index (Table 5), i.e. its regimes overlap strongly in feature space. Its silhouette is significantly lower than that of K-means (Holm $p = 0.0004$), PCA + K-means (0.00003) and the autoencoder + K-means (0.00002), and not different from the HMM (Holm $p = 0.053$). Regimes are more persistent than K-means clusters (dwell 10.2 vs 4.7 days, $p = 0.003$) and shorter than HMM regimes (12.7 days, $p = 0.030$), whose long dwell reflects one state occupying most of the record in several patients. Lower silhouette is partly a consequence of temporal smoothing, but without a predictive benefit it cannot be justified by one.

5.6 Stability across seeds and noise

Across ten seeds the median within-patient SD of the training log-likelihood per step was 0.0035 nats, so the attained likelihood is stable; the partitions are less so (mean pairwise ARI 0.88, median 0.97, minimum 0.37 across patients), indicating near-equivalent local optima in some patients. Agreement with the unperturbed labels falls quickly under noise (Table 6); day-level labels are therefore not precise or reproducible, and only patient- and cohort-level patterns are supported.

Table 6 Sensitivity of MS-VAR(3) labels to additive Gaussian noise (standardised units; mean over 20 patients; ARI against the unperturbed fit of the same model).

Noise SD	0.00	0.05	0.10	0.15	0.20	0.30
ARI	1.00	0.78	0.61	0.53	0.41	0.36
Transitions	17.7	18.6	20.4	21.8	24.3	23.7

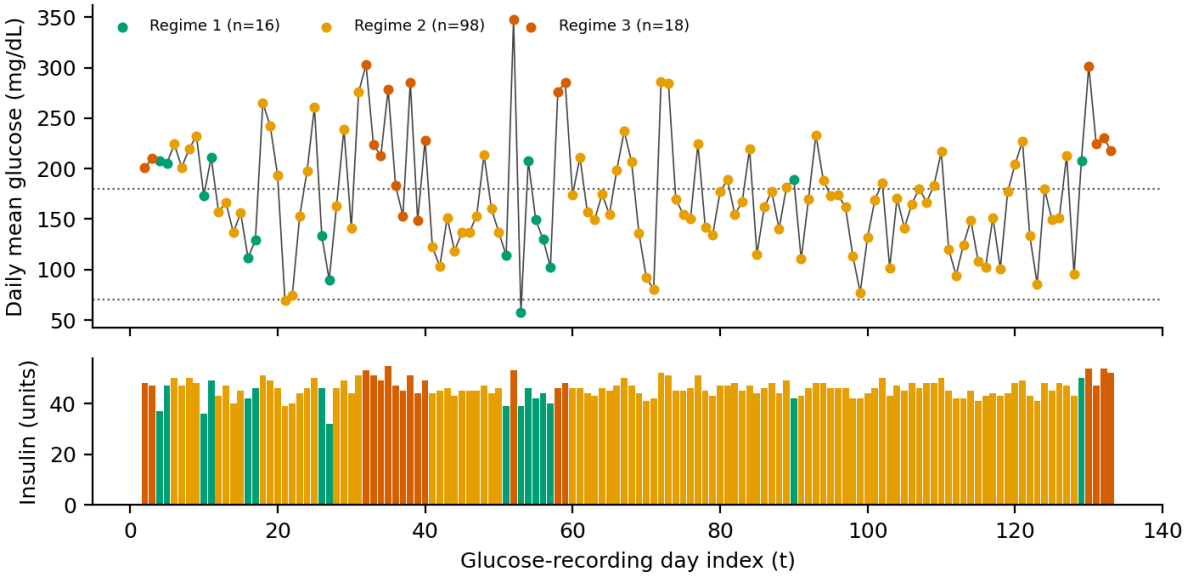


Figure 2 Daily mean glucose (top) and insulin dose (bottom) for patient 20 (133 modelled days), coloured by MS-VAR(3) regime ranked by mean glucose; dotted lines: 70 and 180 mg/dL. The patient was chosen for illustration before the analysis was repeated. The regime sequence is fragmented at the day level, in line with the noise sensitivity; numbering is a post hoc rank, not a disease stage.

5.7 Descriptive glycaemic profile

Regimes of each patient's MS-VAR(3) full-series fit were ranked by mean glucose (all 20 patients populate three regimes; Table 7, Fig. 2). Mean glucose increases from rank 1 to 3 by construction; the share of days with maximum above 180 mg/dL increases and the share with minimum below 70 mg/dL decreases. This association is largely mechanical, since the daily minimum, maximum and mean come from the same readings, so it is a plausibility check, not independent validation. The hypoglycaemic-range share is highest in the lowest-glucose regime, compatible with more intensive dosing, but this reading is post hoc. Mean insulin dose has a large between-patient spread.

Table 7 Descriptive profile of MS-VAR(3) regimes ranked by mean glucose (mean ± SD over patients). Thresholds are operational values, not a diagnosis.

Regime	Days	Glucose (mg/dL)	% days min < 70	% days max > 180	Insulin dose
lowest	29.1	136.8 ± 28.3	40.7 ± 23.1	49.5 ± 29.5	61.8 ± 98.3
middle	54.0	159.5 ± 36.8	31.4 ± 26.9	60.4 ± 29.1	35.3 ± 18.4
highest	24.4	193.5 ± 52.3	17.7 ± 20.1	82.0 ± 29.6	52.8 ± 75.9

6 Discussion and limitations

On this data set a switching autoregression does not predict held-out days better than a single-regime VAR(1): the paired difference is close to zero with a confidence interval containing both signs, BIC selects one or two regimes in most patients, and the full model over-fits. The advantage over an i.i.d.-emission HMM in the six-feature specification is not a property of regime switching: a one-regime VAR(1) reproduces it, it arises from an exact lagged identity, and it disappears when the redundant features are removed. In the main specification autoregression helps against i.i.d. models only through autocorrelated moving averages; with raw features it adds nothing. Daily glucose and insulin thus carry little predictable day-to-day structure beyond slowly varying levels, and short series limit what can be estimated. The fitted regimes are persistent, ordered by glucose and populated by all patients, so they are a usable descriptive segmentation of a patient's record; they are not shown to be physiological states.

Four design choices proved decisive: a single-regime baseline with the same dynamics; screening features for exact linear relations to the lagged state; normalising held-out likelihood per observation; and multiplicity corrections within declared families. Because the observation operator was the identity, the study does not test whether a continuous latent physiological state would help; that needs data such as continuous glucose monitoring with meal and exercise annotations.

Limitations. The cohort is small (20 of 70) and selected by recording length; the data are a single historical sample without demographic or diagnostic metadata, and no diabetes type is claimed. Held-out evaluation uses one temporal split with 12–33 transitions per patient. The main feature set was chosen from the identity diagnostic rather than from the held-out comparison, and results for all three sets are reported. λ, ε and $M = 3$ were fixed a priori; EM with ridge is not guaranteed monotone. Logged insulin is not total exposure and plausibility of glucose values was not screened. The glycaemic profile is post hoc and partly mechanical, with no clinical validation. All models were implemented from scratch (NumPy/SciPy) and checked by likelihood and valid-probability tests, but not against an independent library.

7 Conclusion

A Markov-switching vector autoregression was fitted to real, unlabelled self-monitoring records from the AIM'94 data under a documented pipeline, and its out-of-sample value was assessed against baselines isolating persistence, autoregression and regime-specific dynamics. The result is negative for prediction: regime switching did not improve held-out prediction over a single-regime VAR(1), model selection did not support three regimes, and an apparent advantage over an HMM in a six-feature specification was traced to an exact linear identity among features. The regimes remain useful as a patient-level descriptive segmentation ordered by glucose, but day-level labels are noise-sensitive and not validated as physiological states. Further work should use denser continuous-monitoring data, rolling-origin evaluation, model-selected regime counts and clinician-assigned labels.

Declarations

Conflicts of interest. The authors declare no conflict of interest.

Data availability. The AIM'94 diabetes dataset is publicly available in the UCI Machine Learning Repository; the analysis code and per-patient result tables are available from the authors on reasonable request.

References

[1] Rabiner L. R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*. 77(2): 257–286. doi: https://doi.org/10.1109/5.18626.

[2] Hamilton J. D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*. 57(2): 357–384.

[3] Krolzig H.-M. 1997. *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*. Berlin: Springer. 357 p.

[4] Ghahramani Z., Hinton G. E. 2000. Variational learning for switching state-space models. *Neural Computation*. 12(4): 831–864. doi: https://doi.org/10.1162/089976600300015619.

[5] Cobelli C., Dalla Man C., Sparacino G., Magni L., De Nicolao G., Kovatchev B. 2009. Diabetes: models, signals, and control. *IEEE Reviews in Biomedical Engineering*. 2: 54–96. doi: https://doi.org/10.1109/RBME.2009.2036073.

[6] Bergman R. N., Ider Y. Z., Bowden C. R., Cobelli C. 1979. Quantitative estimation of insulin sensitivity. *American Journal of Physiology*. 236(6): E667–E677. doi: https://doi.org/10.1152/ajpendo.1979.236.6.E667.

[7] Weir G. C., Bonner-Weir S. 2004. Five stages of evolving beta-cell dysfunction during progression to diabetes. *Diabetes*. 53(Suppl. 3): S16–S21. doi: https://doi.org/10.2337/diabetes.53.suppl_3.s16.

[8] Kalman R. E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*. 82(1): 35–45. doi: https://doi.org/10.1115/1.3662552.

[9] Wang Q., Molenaar P., Harsh S. [et al.] 2014. Personalized state-space modeling of glucose dynamics for type 1 diabetes using continuously monitored glucose, insulin dose, and meal intake: an extended Kalman filter approach. *Journal of Diabetes Science and Technology*. 8(2): 331–345. doi: https://doi.org/10.1177/1932296814524080.

[10] Clausen H. G., Leth J., Knudsen T., Schiøler H. 2020. State estimation in type 2 diabetes using the continuous-discrete unscented Kalman filter. *IFAC-PapersOnLine*. 53(2): 16500–16505. doi: https://doi.org/10.1016/j.ifacol.2020.12.754.

[11] Dempster A. P., Laird N. M., Rubin D. B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*. 39(1): 1–38.

[12] Billio M., Monfort A. 1998. Switching state-space models: likelihood function, filtering and smoothing. *Journal of Statistical Planning and Inference*. 68(1): 65–103. doi: https://doi.org/10.1016/S0378-3758(97)00136-5.

[13] Kahn M. 1994. Diabetes [data set]. *UCI Machine Learning Repository*. URL: https://archive.ics.uci.edu/dataset/34/diabetes.

[14] Pedregosa F., Varoquaux G., Gramfort A. [et al.] 2011. Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*. 12: 2825–2830.

[15] Hubert L., Arabie P. 1985. Comparing partitions. *Journal of Classification*. 2: 193–218.

[16] Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. 6(2): 65–70.

[17] Schwarz G. 1978. Estimating the dimension of a model. *Annals of Statistics*. 6(2): 461–464.

[18] Rousseeuw P. J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*. 20: 53–65.

[19] Caliński T., Harabasz J. 1974. A dendrite method for cluster analysis. *Communications in Statistics*. 3(1): 1–27.

[20] Davies D. L., Bouldin D. W. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. PAMI-1(2): 224–227.

УДК 519.246.8:616.379-008.64

МАРКОВСКАЯ ПЕРЕКЛЮЧАЮЩАЯСЯ ВЕКТОРНАЯ АВТОРЕГРЕССИЯ ДЛЯ СКРЫТЫХ ГЛИКЕМИЧЕСКИХ РЕЖИМОВ В РЕАЛЬНЫХ ДАННЫХ САМОКОНТРОЛЯ ДИАБЕТА: ОЦЕНКА НА ОТЛОЖЕННОЙ ВЫБОРКЕ

[1] ***Исомиддинов А.И.,*** [1*] ***Акбарова А.А.,*** [2] ***Акбаров Б.Х.***

[*] akbarovaaa1991@gmail.com

[1]Наманганский государственный технический университет, 160115, Узбекистан, г. Наманган, ул. Косонсой, 12;

[2]Международный университет Кимё в Ташкенте, 100121, Узбекистан, г. Ташкент, ул. Усмана Насыра, 156

Записи самоконтроля глюкозы и доз инсулина образуют нерегулярные временные ряды, физиологический режим которых не наблюдается. В работе к необработанным журналам событий открытого набора UCI AIM'94 применяется марковская переключающаяся векторная авторегрессия MS-VAR(1) со своими для каждого режима свободными членами, матрицами авторегрессии и ковариациями, и исследуется, что переключение режимов добавляет к прогнозу на отложенной выборке. Модель оценивается для каждого пациента отдельно EM-алгоритмом с точным сглаживанием «вперёд–назад» без использования меток режимов; по априорным критериям отобраны 20 пациентов из 70 (2189 дней с измерениями глюкозы). Модели сравнивались по прогностическому логарифмическому правдоподобию на один шаг вперёд в расчёте на день на последних 20 % ряда с однорежимной VAR(1), гауссовской скрытой марковской моделью (CMM), однопроходной переключающейся моделью и статической гауссовской моделью с помощью парных критериев Уилкоксона с поправкой Холма и BIC. Набор признаков, содержащий одновременно уровни и первые разности, вырожден для авторегрессии первого порядка: две компоненты являются точными линейными функциями предыдущего состояния, и наименьшее собственное значение ковариации остатков совпало с порогом регуляризации у всех 20 пациентов. В этом варианте преимущество над CMM воспроизводится, но однорежимная VAR(1) показывает тот же результат. После исключения избыточных признаков MS-VAR(3) не лучше однорежимной VAR(1) (медиана парной разности $-0{,}31$ нат/день, 95 %-й ДИ $[-1{,}11;\ 0{,}89]$), BIC выбирает не более двух режимов у 17 из 20 пациентов, а преимущество над CMM ($+2{,}33$ нат/день) не сохраняется после поправки Холма ($p = 0{,}069$). Выделенные режимы устойчивы во времени (среднее время пребывания 10,2 дня) и упорядочены по средней глюкозе (137, 160, 194 мг/дл), однако метки отдельных дней чувствительны к шуму (ARI 0,78 при СКО шума 0,05). Таким образом, переключение режимов оправдано как описательная сегментация, но не как улучшение прогноза.

Ключевые слова: данные самоконтроля диабета, марковская переключающаяся векторная авторегрессия, скрытая марковская модель, скрытый режим, оценка на отложенной выборке, правдоподобие, выбор модели

Цитирование: *Исомиддинов А.И., Акбарова А.А., Акбаров Б.Х.* Марковская переключающаяся векторная авторегрессия для скрытых гликемических режимов в реальных данных самоконтроля диабета: оценка на отложенной выборке // Проблемы вычислительной и прикладной математики. – 2026. – № 5(76). – С. 60-71.

DOI: 10.71310/pcam.5_76.2026.05

№ 5(76) 2026

ISSN 2181-8460

HISOBLASH VA AMALIY MATEMATIKA MUAMMOLARI

ПРОБЛЕМЫ ВЫЧИСЛИТЕЛЬНОЙ И ПРИКЛАДНОЙ МАТЕМАТИКИ

PROBLEMS OF COMPUTATIONAL AND APPLIED MATHEMATICS

AiRI
RAQAMLI TEXNOLOGIYALAR VA SUN'IY INTELLEKTNI RIVOJLANTIRISH ILMIY-TADQIQOT INSTITUTI

ПРОБЛЕМЫ ВЫЧИСЛИТЕЛЬНОЙ И ПРИКЛАДНОЙ МАТЕМАТИКИ

№ 5(76) 2026

Журнал основан в 2015 году.
Издается 6 раз в год.

Учредитель:
Научно-исследовательский институт развития цифровых технологий и искусственного интеллекта.

Главный редактор:
Равшанов Н.

Заместители главного редактора:
Арипов М.М., Шадиметов Х.М., Ахмедов Д.Д.

Ответственный секретарь:
Убайдуллаев М.Ш.

Редакционный совет:
Азамов А.А., Алоев Р.Д., Амиргалиев Е.Н. (Казахстан), Арушанов М.Л., Абдуллозода С.Ф. (Таджикистан), Боборахимов Б.И., Бурнашев В.Ф., Джумаёзов У.З., Загребина С.А. (Россия), Задорин А.И. (Россия), Игнатьев Н.А., Ильин В.П. (Россия), Иманкулов Т.С. (Казахстан), Исмагилов И.И. (Россия), Кабанихин С.И. (Россия), Курбонов Н.М., Маматов Н.С., Мирзаев Н.М., Мурадов Ф.А., Назирова Э.Ш., Нормуродов Ч.Б., Нуралиев Ф.А., Нуралиев Ф.М., Опанасенко В.Н. (Украина), Расулмухамедов М.М., Садуллаева Ш.А., Старовойтов В.В. (Беларусь), Хаётов А.Р., Халджигитов А., Хамдамов Р.Х., Хужаев И.К., Хужаеров Б.Х., Эшмаматова Д.Б., Дустмуродова Ш.Ж., Чье Ен Ун (Россия), Шабозов М.Ш. (Таджикистан), Dimov I. (Болгария), Li Y. (США), Mascagni M. (США), Min A. (Германия), Singh M. (Южная Корея).

Журнал зарегистрирован в Агентстве информации и массовых коммуникаций при Администрации Президента Республики Узбекистан.
Свидетельство №0856 от 5 августа 2015 года.

ISSN 2181-8460, eISSN 2181-046X

При перепечатке материалов ссылка на журнал обязательна.
За точность фактов и достоверность информации ответственность несут авторы.

Адрес редакции:
100125, г. Ташкент, м-в. Буз-2, 17А.
Тел.: +(998) 71 263-41-98.
Э-почта: journals@airi.uz.
Веб-сайт: https://journals.airi.uz.

Дизайн и вёрстка:
Шарипов Х.Д.

Отпечатано в типографии НИИ РЦТИИ.
Подписано в печать 28.10.2026 г.
Формат 60x84 1/8. Заказ № 6. Тираж 100 экз.

PROBLEMS OF COMPUTATIONAL AND APPLIED MATHEMATICS

No. 5(76) 2026

The journal was established in 2015.
6 issues are published per year.

Founder:
Digital Technologies and Artificial Intelligence Development Research Institute.

Editor-in-Chief:
Ravshanov N.

Deputy Editors:
Aripov M.M., Shadimetov Kh.M., Akhmedov D.D.

Executive Secretary:
Ubaydullaev M.Sh.

Editorial Council:
Azamov A.A., Aloev R.D., Amirgaliev E.N. (Kazakhstan), Arushanov M.L., Abdullozoda S.F. (Tajikistan), Boborakhimov B.I., Burnashev V.F., Djumayozov U.Z., Zagrebina S.A. (Russia), Zadorin A.I. (Russia), Ignatiev N.A., Ilyin V.P. (Russia), Imankulov T.S. (Kazakhstan), Ismagilov I.I. (Russia), Kabanikhin S.I. (Russia), Kurbonov N.M., Mamatov N.S., Mirzaev N.M., Muradov F.A., Nazirova E.Sh., Normurodov Ch.B., Nuraliev F.A., Nuraliev F.M., Opanasenko V.N. (Ukraine), Sadullaeva Sh.A., Starovoitov V.V. (Belarus), Khayotov A.R., Khaldjigitov A., Khamdamov R.Kh., Khujaev I.K., Khujayorov B.Kh., Eshmamatova D.B., Dustmurodova Sh.J., Chye En Un (Russia), Shabozov M.Sh. (Tajikistan), Dimov I. (Bulgaria), Li Y. (USA), Mascagni M. (USA), Min A. (Germany), Singh M. (South Korea).

The journal is registered by Agency of Information and Mass Communications under the Administration of the President of the Republic of Uzbekistan.
Certificate of Registration No. 0856 of 5 August 2015.

ISSN 2181-8460, eISSN 2181-046X

At a reprint of materials the reference to the journal is obligatory.
Authors are responsible for the accuracy of the facts and reliability of the information.

Address:
100125, Tashkent, Buz-2, 17A.
Tel.: +(998) 71 263-41-98.
E-mail: journals@airi.uz.
Web-site: https://journals.airi.uz.

Layout design:
Sharipov Kh.D.

DTAIRI printing office.
Signed for print 28.10.2026
Format 60x84 1/8. Order No. 6. Print run of 100 copies.

Содержание

Расулов А.С., Раимова Г.М., Рахматов М.Й.
Метод Монте-Карло для решения нелинейного кинетического уравнения Больцмана 8

Эшкулов М.У., Хамдамов Р.Х.
Многокритериальная оптимизация системы подъёма воды с использованием солнечных фотоэлектрических панелей 22

Абдуллаев К.Д.
Принцип разделения классов и зависимость результатов кластеризации от меры расстояния 33

Рамазонов Ш.Ш.
Селекция объектов по отношениям с ближайшими объектами противоположного класса и сохранение локальной структуры обучающей выборки . . 46

Исомиддинов А.И., Акбарова А.А., Акбаров Б.Х.
Марковская переключающаяся векторная авторегрессия для скрытых гликемических режимов в реальных данных самоконтроля диабета: оценка на отложенной выборке 60

Ким К.К.
Сравнительный анализ XAI-подходов для обнаружения синтетической речи 72

Эшмаматова Д.Б., Солижанова К.А.
Пфаффианова структура и спектральная динамика дискретного оператора Лотки–Вольтерры: подход к интеграции геологических данных 95

Джумаёзов У.З., Абдирахмонова М.Н., Усмонов Л.С.
Численное решение трёхмерной связанной динамической задачи термоупругости в деформациях 119

Бакаев И.И., Ахмеджанова Д.А.
α-справедливая оптимизационная модель и численные алгоритмы выбора общей лексической нормы тюркских языков 137

Равшанов Н., Камолиддинова Ф., Lu Y.
Моделирование переноса и диффузии загрязняющих веществ в пограничном слое атмосферы с учётом орографии местности и метеорологических факторов 148

Contents

Rasulov A.S., Raimova G.M., Rakhmatov M.Y.
Monte Carlo solution of the nonlinear kinetic Boltzmann equation 8
Eshkulov M.U., Khamdamov R.Kh.
Multi-criteria optimization of solar photovoltaic-based water lifting systems . . . 22
Abdullaev K.D.
The principle of class separation and the dependence of clustering results on the distance measure 33
Ramazonov Sh.Sh.
Instance selection based on nearest-enemy relations and preservation of the local structure of a training sample 46
Isomiddinov A.I., Akbarova A.A., Akbarov B.X.
Markov-switching vector autoregression for latent glycaemic regimes in real-world diabetes self-monitoring data: an out-of-sample evaluation 60
Kim K.K.
Comparative analysis of XAI approaches for synthetic speech detection 72
Eshmamatova D.B., Solijanova K.A.
Pfaffian structure and spectral dynamics of a discrete Lotka–Volterra operator: a framework for geological evidence integration 95
Djumayozov U.Z., Abdirakhmonova M.N., Usmonov L.S.
Numerical solution of a three-dimensional coupled dynamic thermoelasticity problem in terms of strains 119
Bakayev I.I., Akhmedjanova D.A.
An α-fair optimization model and numerical algorithms for selecting a common lexical norm of Turkic languages 137
Ravshanov N., Kamoliddinova F., Lu Y.
Modeling pollutant transport and diffusion in the atmospheric boundary layer considering terrain orography and meteorological factors 148