

パーソナライズされた打鍵ダイナミクスによる認知状態推定モデル

Gamita Novelist Development Team

2026年3月29日

1 概要

本アルゴリズムは、打鍵パターンのみから執筆者の集中強度を連続値スコア $F_t \in [0, 100]$ として推定するリアルタイムエンジンである。

本稿は、Gamita Novelist で研究している打鍵ダイナミクスによる認知状態推定の初期研究成果を公開するものである。本稿の公開の目的は、Gamita Novelist が創作活動支援と向き合っている過程を、利用ユーザーに少しでも知ってもらうためである。正式な学術論文や先行研究発表などとして発表するものではない。

ここで定式化されたアルゴリズムは、改良されたのちに Gamita Novelist のテキストエディタに応用されている。現行の実装では本稿の手法を礎として独自のアクティビティ機能へと発展し、集中力のみならずストレス状態のリアルタイム推定にも対応している。

設計上の核となる原則は以下の通り。

1. **絶対量ではなく個人内相対量**を測定する。同じ打鍵速度でも、ある人にとってはフローであり、別の人にとっては低活動である。
2. **沈黙 ≠ 非集中**として扱う。小説家は文章を考える間、止まる。思考的休止を寛容に受け入れる時定数はユーザーごとに異なり、アルゴリズムが**自動学習**する。
3. **固有値を排除**する。アルゴリズム動作を左右する主要な時定数は実行時に統計量から自動算出する。開発者の直感による「感度調整」を不要にする。
4. **長期記憶と短期応答の両立**を図る。

5. 堅牢性を絶対に保証する。NaN・Inf・負の分散は生成しない。

本論文がもたらす進歩性

既存のタイピング監視システム等は「5 秒間入力がなければ非集中」といった固定値に依存しており、プロの長考や個人の生体的なリズムを誤判定する構造的な課題があった。また、そのような一律的な判定モデルに由来する過信・誤信により信頼性を損ねている現状がある。本研究のコアとなる貢献は以下の 3 点である。

1. **生体リズムの動的写像**：執筆者固有の「思考のための必要な休止」と「活動の連鎖」をリアルタイムで逐次学習し、これを単なる統計量ではなく、「信号処理フィルタの時定数 (τ) へ自己適応的に写像する」という独自の制御モデルを新たに構築した。
2. **長期忘却と再学習モデリング**：怪我や加齢によるスランプといった長期的なライフスタイルの変化（恒久的なコンディションの低下）に対し、「過去のベースラインによる誤判定」を数理的に解放し、再適応を促すための忘却関数を定式化した。
3. **統計的堅牢性の検証**：固定値ベースラインとの比較において、本手法が特異な打鍵パターン（長考型や早筆型など）においても個人の正常なスコアを保ちつつ、微小な過集中や散漫を統計的に有意な水準 ($p < 0.001$) で分離可能であることをシミュレーション上で示した。

シミュレーション設計の妥当性について

■ **ペルソナ設計の根拠** 打鍵速度レンジの上限は REALFORCE TYPING CHAMPIONSHIP 2025 出場選手の記録（約 1,000 kpm \approx 16.7 keys/s）を参考に設定し、下限は一般的な PC 操作が正常に行える最低速度（約 3 keys/s）を基準とした。各ペルソナの速度パラメータはこのレンジ内に収まる設計となっている。

8 種類のスタイルは「多様性の網羅」ではなく、本アルゴリズムの設計原則それぞれに対するストレステストとして選定した。すなわち、P04・P08（長考型）は「沈黙 \neq 非集中」原則、P05（早筆型）は τ_{rise} 上限クランプ機能、P01（気まぐれ型）はエイジング機能、P06（プロ規則型）は τ_{fall} の精度、P08（プロ長考型）はプロ打鍵密度と長考

休止の共存という、それぞれが独立した設計上の課題を代表するケースとなっている。

本検証はキーストロークの先行研究を検証・比較するものではなく、ユーザーが実際に使うツール・手法との比較を目的としている。自己啓発やビジネス書などで紹介されている手法を検証の対象とした。これにより、一般的に有効とされているモチベーション維持や自己認知が役立つのかをランダム性が高いシミュレーションを行うことで検証する。

本アルゴリズムを応用した製品版には、学習データの一部に提供された実ユーザーのデータを使用して追加検証を行い、精度の向上を図っている。

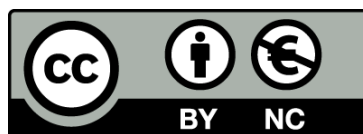
■比較対象の選定理由 比較対象 (Binary、Linear、EMA、Persistence、Static Z、ポモドーロ各種) はすべて、実際のユーザーが代替手段として選択しうる現実的なツール・手法および統計学的な手法から選定した。

■効果量の解釈 Cohen's d が高値 (例: 2 以上) を示すのは、シミュレーション上で「絶好調状態」と「散漫状態」の条件差が設計者によって制御されているためであり、これはシミュレーション固有の特性である。独立した実データによる検証では、効果量は一般に縮小する。本稿に記載の数値は、アルゴリズムの設計特性をシミュレーション条件下で評価した指標として解釈すべきものであり、実環境での絶対的な性能を保証するものではない。

ライセンス表記

本稿および本稿に記載のアルゴリズムは、クリエイティブ・コモンズ表示 - 非営利 4.0 国際 (CC BY-NC 4.0) ライセンスの下に提供されている。詳細については、以下に示す URL を参照されたい。

<https://www.gamita.net>



個別に商用利用を希望する場合は、上記 URL より問い合わせられたい。

2 入力と前提

各ティックの入力：

- t ：現在のタイムスタンプ（秒、単調増加）
- K_t ：前回ティック以降の打鍵数（非負整数）

セッション継続関数：

$$S(t_{\text{prev}}, t) = \begin{cases} \text{true} & t > t_{\text{prev}} \wedge \Delta t_{\text{raw}} \leq \Delta t_{\text{max}} \\ \text{false} & \text{otherwise (起動直後・スリープ復帰等)} \end{cases} \quad (1)$$

3 状態変数と初期条件

変数	意味	型	初期値	永続
t_{prev}	前回タイムスタンプ	float	未定義	No
λ	レート移動平均	float	0	No
λ^{slow}	閾値学習用低速平均	float	0	Yes
\bar{X}	密度移動平均	float	0	Yes
μ	対数密度の平均	float	0	Yes
v	対数密度の分散	float	1.0	Yes
t_{total}	加重有効活動時間	float	0	Yes
$t_{\text{last_session}}$	前回セッション終了時刻	float	t_0	Yes
T_{burst}	バースト持続時間の学習値	float	30	Yes
n_{burst}	バーストイベント観測回数	uint	0	Yes
T_{pause}	休止持続時間の学習値	float	20	Yes
n_{pause}	休止イベント観測回数	uint	0	Yes
Δ^{idle}	累積沈黙時間	float	0	No
$\Delta t_{\text{since_resume}}$	再開からの経過時間	float	∞	No
$t_{\text{since_depart}}$	出発からの経過時間	float	∞	No
A_{prev}	前回活動度	float	0	No
S	平滑化スコア	float	50	No
F	最終表示スコア	float	50	No

$v_0 = 1.0$ の理由：初期値を 0 にすると $\sigma \approx \sqrt{\epsilon}$ となり、最初の数ティックで Z 値が異常に大きくなる。 $v_0 = 1.0$ は「対数密度は標準偏差 ± 1 程度のばらつき」という穏やかな事前仮定であり、信頼度 γ_t と組み合わせて初期安定化を担う。

$T_{burst,0} = 30s, T_{pause,0} = 20s$ の理由：学習前の安全なデフォルト。これらは数回の活動サイクルで実測値に収束する。実測が 1 回蓄積されれば既に部分的に自分に合った値となる。

4 起動時フック（パイプライン外）

メインループ開始前に一度だけ実行する。

$$t_{total} \leftarrow t_{total} \cdot \exp\left(-\frac{t_{now} - t_{last_session}}{\tau_{decay}}\right) \quad (2)$$

前回セッションからの経過時間に応じて t_{total} を減衰させる。長期離脱後は統計感度が部分的に復活し、再開後の再較正が速くなる。

5 自動算出される時定数の設計

本節では、従来の固定値パラメータを自動算出に置き換える核心部分を説明する。

5.1 動機：固定値の問題

τ_{rise} や τ_{fall} を固定した場合、以下の問題が生じた。

- 速いタイポリズム（10 秒バースト）の人には過敏または過鈍な反応
- 遅いタイポリズム（3 分バースト）の人には不適切な反応速度
- チューニングに開発者の主観が混入する

5.2 解決策：バースト・休止時間の観測学習（ハイブリッド Welford→EMA）

アルゴリズムは活動状態の遷移を観測し、各ユーザーの「呼吸のリズム」を学習する。

- バースト持続時間 T_{burst} ：
活動中 ($A_t \geq 0.5$) が続く時間の学習値。「出発イベント」(A が 0.5 を下回る瞬

間)に更新される。

- 休止持続時間 T_{pause} :

非活動 ($A_t < 0.5$) が続く時間の学習値。「再開イベント」(A が 0.5 を上回る瞬間)に更新される。

■更新方式：ハイブリッド Welford→EMA 最初の N_{switch} イベントは Welford の逐次平均 ($1/n$ 更新) を使用し、以降は EMA (α_{event} 更新) に切り替える。

$$\alpha^*(n) = \begin{cases} 1/n & n \leq N_{switch} \\ \alpha_{event} & n > N_{switch} \end{cases} \quad (3)$$

$$n \leftarrow n + 1 \quad (4)$$

$$T \leftarrow T + \alpha^*(n) \cdot (x_{new} - T) \quad (5)$$

($T \in \{T_{burst}, T_{pause}\}$ 、 x_{new} は今回の観測値、 $n \in \{n_{burst}, n_{pause}\}$)

設計根拠：

- Welford フェーズ ($n \leq N_{switch}$)：初期値 T_0 (デフォルト値) の影響が n 回で $1/n$ に薄まり、素直に真の平均へ収束する。
- EMA フェーズ ($n > N_{switch}$)：学習率が α_{event} で固定されるため、ユーザーのリズムが変化しても追従し続ける。純 Welford では $n \rightarrow \infty$ で更新量がゼロになり、長期の変化に追従できなくなるためこの切り替えが必要。

5.3 適応的時定数の算出式

■上昇時定数 $\tau_{rise,t}$ ：バーストの典型的な長さに比例して、速く検出する。

$$\tau_{rise,t} = \text{clip}\left(\frac{T_{burst}}{20}, \tau_{rise,min}, \tau_{rise,max}\right) \quad (6)$$

係数 $1/20$ の意味：バースト開始から $T_{burst}/20$ 秒で λ は X_t に $1 - e^{-1} \approx 63\%$ 追従する。バーストの 5% 未満の時間で検出が始まる。

例： $T_{burst} = 30\text{s} \Rightarrow \tau_{rise} = 1.5\text{s}$ (3 秒で 86% 追従)

■下降時定数 $\tau_{fall,t}$ ： 通常の休止より少し長い間は、レートを高いまま保つ。

$$\tau_{fall,t} = \text{clip}(T_{pause} \cdot 0.6, \tau_{fall,min}, \tau_{fall,max}) \quad (7)$$

係数 0.6 の意味：典型的な休止時間 T_{pause} の 60% が経過した時点で λ は $e^{-T_{pause}/\tau_{fall}} = e^{-1/0.6} \approx 0.19$ 倍になる。閾値 $\lambda_{th} = 0.3\lambda^{slow}$ に対し、 $\lambda_0 = \lambda^{slow}$ から始めると $0.19 < 0.3$ つまり**通常の休止の 60% 後に非活動状態に入る**。これにより「いつもより少し長い沈黙」でようやく非活動と判定される。

例： $T_{pause} = 20\text{s} \Rightarrow \tau_{fall} = 12\text{s}$ (12 秒休止で 63% 低下)

■シグモイド c_t ： 新しい状態変数なしで算出できる。 σ_t (打鍵密度の対数標準偏差) はすでに統計量として計算されており、ユーザーの打鍵パターンのばらつきを直接反映している。

$$c_t = \text{clip}\left(\frac{\sigma_t}{2}, c_{min}, c_{max}\right) \quad (8)$$

設計根拠： σ_t が大きい (ばらつきの大きいバースト型ユーザー) ほど c_t が大きくなり、活動度の遷移がよりなめらかになる。 σ_t が小さい (安定持続型ユーザー) ほど c_t が小さくなり、活動・非活動の境界がより明確に引かれる。適応的な個人内較正が自動で行われる。

例： $\sigma_t = 1.0$ (中程度のばらつき) $\Rightarrow c_t = 0.5$ (本アルゴリズムのデフォルト値と一致)

5.4 固定値として残るパラメータとその理由

パラメータ	理由
$r_{th} = 0.3$	「活動の定義」は設計思想。データから一意に導けない。
$D_{max} = 8$	安全弁 (上限クランプ)。動的にすると不安定化リスクがある。
$\lambda_{th,min} = 1.0$	コールドスタート時の数値安全弁。
$\tau_{slow}, \tau_{density}$	統計の時間スケール設定。設計的選択。
$T_{floor}^{sil}, T_{ceil}^{sil}$	スコアの語義定義。変えると表示意味が変わる。

6 処理パイプライン

Step 1: 時間制御 ($\Delta t = 0$ なら即 return)

- Step 2:** 打鍵密度 (X_t, Y_t)
- Step 3:** 適応的レート推定 ($\tau_{rise,t}$ と $\tau_{fall,t}$ を自動算出してから λ_t を更新)
- Step 4:** 活動閾値 ($\lambda_t^{slow}, \lambda_{th,t}$)
- Step 5:** 適応的活動判定 (c_t を自動算出してから A_t を計算)
- Step 6:** 遷移イベント処理 (出発 / 再開イベントを判定し T_{burst}, T_{pause}, S を更新)
- Step 7:** 個人内相対密度 (\bar{X}_t, D_t)
- Step 8:** 統計量更新 ($t_{total}, \mu_t, \nu_t, \sigma_t$)
- Step 9:** 信頼度 (γ_t)
- Step 10:** 瞬間強度 (Z_t, T_t)
- Step 11:** 沈黙時間・ターゲット ($\Delta_t^{idle}, T_t^{silence}$)
- Step 12:** ダイナミクス (S_t , グレース適用)
- Step 13:** 最終出力 (F_t)
- Step 14:** 状態保存

7 各ステップ詳細

Step 1: 時間制御

$$\Delta t = \begin{cases} \min(\Delta t_{raw}, \Delta t_{max}) & S = \text{true} \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

$\Delta t = 0$ の場合、以降すべてスキップし return する。

Step 2: 打鍵密度

$$X_t = \frac{K_t}{\Delta t}, \quad Y_t = \ln(X_t + 1) \quad (10)$$

Step 3: 適応的レート推定

まず自動算出された時定数を求める：

$$\tau_{rise,t} = \text{clip}\left(\frac{T_{burst}}{20}, \tau_{rise,min}, \tau_{rise,max}\right) \quad (11)$$

$$\tau_{fall,t} = \text{clip}(T_{pause} \cdot 0.6, \tau_{fall,min}, \tau_{fall,max}) \quad (12)$$

$$\tau_t^{rate} = \begin{cases} \tau_{rise,t} & X_t \geq \lambda_{t-1} \\ \tau_{fall,t} & X_t < \lambda_{t-1} \end{cases} \quad (13)$$

$$\kappa_t^{rate} = 1 - \exp(-\Delta t / \tau_t^{rate}) \quad (14)$$

$$\lambda_t = \lambda_{t-1} + \kappa_t^{rate} (X_t - \lambda_{t-1}) \quad (15)$$

Step 4: 活動閾値

$$\kappa_t^{slow} = 1 - \exp(-\Delta t / \tau_{slow}) \quad (16)$$

$$\lambda_t^{slow} = \lambda_{t-1}^{slow} + \kappa_t^{slow} (\lambda_t - \lambda_{t-1}^{slow}) \quad (17)$$

$$\lambda_{th,t} = \max(r_{th} \cdot \lambda_t^{slow}, \lambda_{th,min}) \quad (18)$$

Step 5: 適応的活動判定

シグモイドを自動算出：

$$c_t = \text{clip}\left(\frac{\sigma_{t-1}}{2}, c_{min}, c_{max}\right) \quad (19)$$

(σ_{t-1} は前ティックの統計量を使用。初回は $\sigma_0 = \sqrt{v_0} = 1.0$)

$$A_t = \frac{1}{1 + \exp\left(-\frac{\lambda_t - \lambda_{th,t}}{c_t \cdot \lambda_{th,t} + \epsilon}\right)} \quad (20)$$

Step 6: 遷移イベント処理

出発イベント（活動 → 非活動）のトリガー：

$$A_{prev} \geq 0.5 \wedge A_t < 0.5 \quad (21)$$

処理 (ハイブリッド Welford → EMA) :

$$\begin{aligned}
 T_{burst}^{obs} &= \Delta t_{since_resume} & (22) \\
 \text{if } T_{\min} \leq T_{burst}^{obs} \leq T_{\max} & \text{ (有効域ガード) :} \\
 n_{burst} &\leftarrow n_{burst} + 1 \\
 \alpha_{burst}^* &= \begin{cases} 1/n_{burst} & n_{burst} \leq N_{switch} \\ \alpha_{event} & \text{otherwise} \end{cases} \\
 T_{burst} &\leftarrow T_{burst} + \alpha_{burst}^* (T_{burst}^{obs} - T_{burst}) \\
 t_{since_depart} &\leftarrow 0 & (23)
 \end{aligned}$$

有効域ガードの意味と設計根拠 :

- $T_{burst}^{obs} < T_{\min}$ (デフォルト 5s) の除外: A_t が 0.5 付近で振動する「チャタリング」により生じる偽の短時間イベントを排除する。
- $T_{burst}^{obs} > T_{\max}$ (デフォルト 600s) の除外: セッション境界とリズムの混同を防ぐ。 T_{burst} と T_{pause} はユーザーの「打鍵のリズム (呼吸)」を学習するものであり、セッション全体の長さ (例: 20 分のまとまった執筆) を学習するものではない。600s (10 分) を超えるバースト・休止は「ワーキングスタイル」であって「リズム」ではないため、ガードにより T_{burst} / T_{pause} の更新対象から除外する。初期化直後の $\Delta t_{since_resume} = \infty$ の除外も兼ねる。

再開イベント (非活動 → 活動) のトリガー :

$$A_{prev} < 0.5 \wedge A_t \geq 0.5 \quad (24)$$

処理 (ハイブリッド Welford → EMA) :

$$T_{pause}^{obs} = t_{since_depart} \quad (25)$$

if $T_{min} \leq T_{pause}^{obs} \leq T_{max}$ (有効域ガード) :

$$n_{pause} \leftarrow n_{pause} + 1$$

$$\alpha_{pause}^* = \begin{cases} 1/n_{pause} & n_{pause} \leq N_{switch} \\ \alpha_{event} & \text{otherwise} \end{cases}$$

$$T_{pause} \leftarrow T_{pause} + \alpha_{pause}^* (T_{pause}^{obs} - T_{pause})$$

$$w = 1 - \exp(-\Delta^{idle} / \tau_{session}) \quad (26)$$

$$S \leftarrow w \cdot 50 + (1 - w) \cdot S \quad (27)$$

$$\Delta^{idle} \leftarrow 0 \quad (28)$$

$$\Delta t_{since_resume} \leftarrow 0 \quad (29)$$

Step 6 が更新した S を、同フレームの Step 12 が S_{t-1} として使用する。

Step 7: 個人内相対密度

$$\kappa_t^{den} = 1 - \exp(-\Delta t / \tau_{density}) \quad (30)$$

$$\bar{X}_t = \bar{X}_{t-1} + \kappa_t^{den} \cdot A_t \cdot (X_t - \bar{X}_{t-1}) \quad (31)$$

$$D_t = \min\left(\frac{X_t + \epsilon}{\bar{X}_{t-1} + \epsilon}, D_{max}\right) \quad (32)$$

Step 8: 統計量更新

$$t_{total} \leftarrow t_{total} \cdot \exp(-\Delta t / \tau_{decay}) + A_t \cdot \Delta t \quad (33)$$

$$\tau_{stat,t} = \tau_{stat,0} + (\tau_{stat,\infty} - \tau_{stat,0}) \times (1 - \exp(-t_{total} / \tau_{stat,grow})) \quad (34)$$

$$\alpha_t = 1 - \exp(-\Delta t / \tau_{stat,t}) \quad (35)$$

$$\mu_t = \mu_{t-1} + \alpha_t \cdot A_t \cdot (Y_t - \mu_{t-1}) \quad (36)$$

$$v_{raw} = v_{t-1} + \alpha_t \cdot A_t \cdot ((Y_t - \mu_{t-1})^2 - v_{t-1}) \quad (37)$$

$$v_t = \max(v_{raw}, v_{min}) \quad (38)$$

$$\sigma_t = \sqrt{v_t} \quad (39)$$

Step 9: 信頼度

$$\gamma_t = 1 - \exp(-t_{total}/\tau_{confidence}) \quad (40)$$

Step 10: 瞬間強度

$$Z_t = \frac{Y_t - \mu_{t-1}}{\sigma_t} \quad (41)$$

$$T_t = 50 + \gamma_t \left(10 \cdot \text{clip}(Z_t, -3, 3) + 8 \cdot \tanh(\ln D_t) \right) \quad (42)$$

Step 11: 沈黙時間とターゲット

$$\Delta_t^{idle} = \Delta_{t-1}^{idle} + (1 - A_t)^2 \cdot \Delta t \quad (43)$$

$$T_t^{silence} = T_{floor}^{sil} + (T_{ceil}^{sil} - T_{floor}^{sil}) \exp(-\Delta_t^{idle}/\tau_{silence}) \quad (44)$$

Step 12: ダイナミクス (グレース付き)

$$g_t = \exp(-\Delta t_{since_resume}/\tau_{grace}) \quad (45)$$

$$\tau_{active} = \tau_{warmup} + g_t \cdot (\tau_{grace_dur} - \tau_{warmup}) \quad (46)$$

$$\tau_t = A_t \cdot \tau_{active} + (1 - A_t) \cdot \tau_{silence} \quad (47)$$

$$\text{Target}_t = A_t \cdot T_t + (1 - A_t) \cdot T_t^{silence} \quad (48)$$

$$S_t = S_{t-1} + (1 - \exp(-\Delta t/\tau_t)) (\text{Target}_t - S_{t-1}) \quad (49)$$

Step 13: 最終出力

$$\beta_t = 1 - \exp(-\Delta t/\tau_{perception}) \quad (50)$$

$$F_t = F_{t-1} + \beta_t (S_t - F_{t-1}) \quad (51)$$

Step 14: 状態保存

$$A_{prev} \leftarrow A_t \quad (52)$$

$$t_{prev} \leftarrow t \quad (53)$$

$$t_{last_session} \leftarrow t \quad (54)$$

$$\Delta t_{since_resume} \leftarrow \Delta t_{since_resume} + \Delta t \quad (55)$$

$$t_{since_depart} \leftarrow t_{since_depart} + \Delta t \quad (56)$$

8 パラメータ表

自動算出されるパラメータ (実行時に毎ティック更新):

パラメータ	算出元
$\tau_{rise,t}$	T_{burst} (学習済みバースト持続時間)
$\tau_{fall,t}$	T_{pause} (学習済み休止持続時間)
c_t	σ_{t-1} (既存の統計量)

固定パラメータ (安全弁・設計上の選択):

パラメータ	意味	値	単位
Δt_{tick}	ティック間隔	1	s
Δt_{max}	最大継続間隔	10	s
$\tau_{rise,min}$	上昇定数の下限	0.2	s
$\tau_{rise,max}$	上昇定数の上限	5.0	s
$\tau_{fall,min}$	下降定数の下限	3.0	s
$\tau_{fall,max}$	下降定数の上限	60.0	s
c_{min}	シグモイド下限	0.1	—
c_{max}	シグモイド上限	2.0	—
N_{switch}	Welford→EMA 切り替え閾値	20	events
α_{event}	EMA フェーズ学習率	0.15	—
τ_{slow}	閾値学習定数	300	s
$\tau_{density}$	密度基準定数	30	s
τ_{warmup}	活動時追従定数	5	s
τ_{grace_dur}	再開グレース期間	30	s
τ_{grace}	グレース減衰定数	15	s
$\tau_{silence}$	非活動時減衰定数	10	s
$\tau_{session}$	セッション冷却定数	180	s
$\tau_{perception}$	表示平滑化定数	10	s
$\tau_{stat,0}$	統計初期定数	50	s
$\tau_{stat,\infty}$	統計長期定数	86400	s
$\tau_{stat,grow}$	統計成長速度	3600	s
$\tau_{confidence}$	ウォームアップ収束定数	90	s
τ_{decay}	エイジング減衰定数	432000	s
r_{th}	閾値比率	0.30	—
$\lambda_{th,min}$	最小閾値	1.0	keys/s
v_{min}	分散下限ガード	0.01	—
D_{max}	密度比上限	8	—
T_{min}	観測有効域の下限 (チャタリング防止)	5	s
T_{max}	観測有効域の上限 (リズム/セッション分離)	600	s
T_{ceil}^{sil}	停止初期目標	40	pt
T_{floor}^{sil}	停止長期目標	20	pt
ϵ	数値安定化定数	10^{-6}	—

9 受け入れテスト仕様 (QA)

以下の7シナリオをパスすることを本番リリースの条件とする。

1. T01: コールドスタート保護

永続化データなし、10 keys/s 連続打鍵。最初の30秒は $F_t \in [45, 55]$ を維持し、90秒後に $F_t > 58$ であること。

2. T02: 思考的休止トレランス (適応後)

少なくとも5回の活動/休止サイクルを経た後 (T_{burst}, T_{pause} 学習済み)、典型的な休止時間と同程度の停止中に $F_t \geq 50$ を維持すること。

3. T03: 長期休止からの復帰

3日以上 of 長期離脱 ($\Delta t_{since_depart} \geq 259200s$) 後、エイジング機能 (τ_{decay}) により t_{total} が減衰すること。再開直後の数分間で新しい打鍵リズムへの再適応が開始され、不自然な低スコアに固定されないこと。

4. T04: バースト感度の自動最適化

早筆型ユーザー ($T_{burst} > 100s$) において、上昇時定数 τ_{rise} が上限値 (5.0s) にクランプされ、一時的な打鍵リズムの揺らぎに対してスコアが過敏に反応 (乱高下) しないこと。

5. T05: 思考的休止の許容判定

学習済みの典型的な休止時間 T_{pause} の60%程度の沈黙が発生しても、活動度 $A_t \geq 0.5$ を維持し、ユーザーの「集中した思考」が途切れたと判定されないこと。

6. T06: 信頼度 (γ_t) の収束性

累積活動時間 t_{total} の増加に従って γ_t が0から1へと単調増加し、約540秒 ($6\tau_{confidence}$) 以内に0.99以上に到達してスコアが安定すること。

7. T07: チャタリング排除性能

5秒未満 (T_{min} 以下) の極めて短い打鍵および停止の繰り返しが発生しても、これらが T_{burst}, T_{pause} の学習値 (イベント平均) に算入されず、リズム学習が汚染されないこと。

状態変数	意味	内容
λ^{slow}	活動閾値の基準	そのユーザーの「定常打鍵速度」
\bar{X}	密度の基準	活動中の「平均的な打鍵の速さ」
μ	統計ベースライン（平均）	対数密度の長期的な中央値
ν	統計ベースライン（分散）	打鍵リズムのばらつき度合い
t_{total}	統計成熟度の指標	累積有効活動時間（エイジング減衰あり）
T_{burst}	上昇・下降感度の基準	「一気に打つ」バーストの典型的な長さ
T_{pause}	上昇・下降感度の基準	「考えて止まる」休止の典型的な長さ

9.1 キャリブレーションフェーズ

本システムのキャリブレーション（校正）には複数の時間スケールがある。各フェーズの終了条件と、そのフェーズ中のアルゴリズムの挙動を以下に示す。

フェーズ	終了条件	挙動
Ph.0 コールドスタート	$t_{total} > 0$	$\gamma_t \approx 0, F_t \approx 50$ を維持
Ph.1 統計ウォームアップ	$t_{total} \approx \tau_{confidence} = 90 \text{ s}$	γ_t が 0.63 に到達、偏差が徐々に反映
Ph.2 統計安定化	$t_{total} \approx \tau_{stat, grow} = 3600 \text{ s}$	τ_{stat} が $\tau_{stat, \infty}$ に向けて収束
Ph.3 リズム学習	$n_{burst} \geq N_{switch} = 20$ かつ $n_{pause} \geq N_{switch}$	τ_{rise}/τ_{fall} が Welford を卒業し EMA へ移行

Ph.4 定常状態

すべての上記を充足

全パラメータが自動適応、長期的に安定動作

キャリブレーションの目安 (1日2時間使用の場合) :

- Ph.1 終了 : 初日の最初の数分
- Ph.2 終了 : 初日の1時間
- Ph.3 終了 : $N_{switch} = 20$ 回の活動/休止サイクル、概ね2~5セッション
- Ph.4 到達 : 2~3日以降 (ユーザーのリズムが安定し始める頃)

UI実装の推奨 : Ph.0~1の間 ($\gamma_t < 0.5$ の間) は、UI上で「キャリブレーション中」インジケータを表示することを強く推奨する。この期間にスコアを「集中度のフィードバック」として使用するのには誤解を招く。

10 エンジニア向けパラメータ操作ガイド

10.1 カテゴリ分類

Category A | 自動算出・手動設定禁止 (実行時に毎ティック上書きされる)

パラメータ	算出元	変更の影響
$\tau_{rise,t}$	T_{burst}	手動設定しても次のイベントで上書きされる
$\tau_{fall,t}$	T_{pause}	同上
c_t	σ_{t-1}	同上

Category B | UXチューニング可能 (体感QAで確認しながら調整する)

パラメータ	効果	推奨調整範囲
τ_{warmup}	スコアの上昇速度	3~15 s
$\tau_{silence}$	スコアの下降速度	5~30 s
$\tau_{perception}$	表示の滑らかさ	5~20 s
τ_{grace_dur}	再開後の緩衝期間	15~60 s
T_{ceil}^{sil}	休止開始直後の目標スコア	30~50 pt
T_{floor}^{sil}	長期放置時の目標スコア	10~25 pt

Category C | 学習挙動のチューニング可能 (変更はリリース単位で行う)

パラメータ	効果	推奨調整範囲
N_{switch}	Welford→EMA の切り替え速さ	10~50 events
α_{event}	EMA フェーズの長期適応速度	0.05~0.3
$\tau_{confidence}$	ウォームアップ保護の長さ	60~300 s
$\tau_{stat,\infty}$	統計の長期記憶	10000~300000 s
τ_{decay}	エイジング速度 (長期離脱の影響)	86400~1296000 s
$\tau_{session}$	セッション分離の閾値	60~600 s

Category D | 安全弁・原則変更しない (値変更には深い理解が必要)

パラメータ	意図	変更時のリスク
$\tau_{rise,min/max}$	適応レートの暴走防止	感度が極端化しスコアが不安定になる
$\tau_{fall,min/max}$	同上	同上
T_{min}, T_{max}	チャタリング・リズム/セッション分離ガード	値を下げすぎると正常なリズムも学習不可になる
v_{min}	負の分散防止	NaN が発生する
D_{max}	密度比の上限クランプ	スコアが爆発する
$\lambda_{th,min}$	コールドスタート安全弁	ゼロ除算・誤判定が発生
ϵ	全体の数値安定化	絶対に変更しない

Category E | 活動定義パラメータ（変更は仕様変更を意味する）

パラメータ	意味	変更の影響
r_{th}	活動/非活動の境界比率	スコア全体の感度が変わる。再キャリブレーション必要
τ_{slow}	活動閾値の学習速度	閾値の追従速度が変わる
$\tau_{density}$	密度基準の学習速度	相対密度 D_t の感度が変わる

10.2 変更手順の推奨

Category B・Cのパラメータを変更する場合は、以下の手順を推奨する。

1. 変更対象のパラメータと期待する挙動の変化を文書化する。
2. T01～T07の受け入れテストをすべて再実行する。
3. 既存ユーザーの $t_{total}, T_{burst}, T_{pause}$ などの現行学習データが新パラメータ下でも意味を持つかを確認する。意味が失われる場合（例： τ_{stat} 系の大幅変更）は $t_{total} \leftarrow 0$ のリセットと `algo_version` の更新を行う。

11 ユースケースシミュレーション（長期適応性シミュレーション）

本アルゴリズムの長期適応性・堅牢性を評価するため、執筆者の想定行動様式を研究者がモデル化した 8 種類の合成ペルソナを構築し、大規模シミュレーションを実施した。各ペルソナに対して 180 日間の長期間シミュレーションを実施した。シミュレーションの結果、主要な統計量 (μ, σ) および時定数 (T_{burst}, T_{pause}) は概ね学習開始から 30~60 日程度で定常状態に達することが確認された。180 日間という設定は、極端に活動頻度が低いユーザーも含めた長期的な数値安定性を保証するための十分条件である。なお本シミュレーションは、メモリおよび計算効率化のために SQLite と NumPy を用いた専用エンジンで実装されており、外部設定ファイルによるパラメータ管理により将来の研究においても完全な再現・拡張が可能である。

11.1 ユーザーペルソナの定義（N1 分析ベース）

従来の「打鍵速度の絶対値の差異」だけでなく、執筆熟練度・モチベーションの波・曜日効果・活動時間帯・疲労によるセッション内減衰を加味した、研究者が設計した 8 種類の合成的な執筆スタイルを定義した。

ID	熟練度	スタイル	ペルソナの特徴
P01	趣味	気まぐれ型	週末に思い立った時だけ数時間執筆。気分次第で活動の波が激しい。
P02	趣味	昼夜逆転型	深夜のテンションで一気を書く。休前日に徹夜気味になることが多い。
P03	志望者	安定リズム型	毎日仕事終わりに1時間は必ず書く習慣がついている。
P04	志望者	長考型	休日にまとまった時間を確保し、プロットや構成を練りながら書く。
P05	志望者	早筆型	Web連載向け。勢いで大量に書くが、修正でよく止まる。
P06	プロ	オフィスアワー	平日10時~18時を作業時間とし、週末は休む。極めて規則的。
P07	プロ	昼夜逆転型	締切前に徹夜で追い込むスタイル。昼夜がズレ込みやすい。
P08	プロ	遅筆・長考型	純文学等、緻密な文章をゆっくり紡ぐ。打つ時は的確だが遅め。

11.2 シミュレーション結果（180日間学習後の適応状態）

180日間の学習（Ph.4定常状態）を経た後の各ペルソナのシステム学習結果、および20時間の典型的な執筆セッション（活動中すなわち $A_t \geq 0.5$ の区間）におけるスコアの統計量を以下に示す。

User ID	T_{burst}	τ_{rise}	T_{pause}	τ_{fall}	μ_{final}	σ_{final}	活動中 F_t の平均 $\pm 1\sigma$
P01_Hobby_Erratic	30.69s	1.53s	34.07s	20.44s	0.48	0.71	36.53 \pm 3.73
P02_Hobby_NightOwl	42.53s	2.13s	19.23s	11.54s	0.95	0.83	41.19 \pm 3.85
P03_Aspiring_Steady	55.21s	2.76s	10.30s	6.18s	1.33	0.77	41.69 \pm 4.12
P04_Aspiring_LongThinker	96.08s	4.80s	46.40s	27.84s	0.59	0.88	39.03 \pm 3.54
P05_Aspiring_Fast	127.59s	5.00s*	18.35s	11.01s	1.65	0.83	45.43 \pm 4.27
P06_Pro_OfficeHours	48.56s	2.43s	7.77s	4.66s	1.83	0.91	44.73 \pm 4.54
P07_Pro_NightOwl	82.17s	4.11s	12.55s	7.53s	1.82	0.91	45.81 \pm 4.25
P08_Pro_Deliberate	56.81s	2.84s	27.85s	16.71s	0.89	0.89	40.12 \pm 4.31

* τ_{rise} は $\tau_{rise,max} = 5.0$ の上限クランプが発動

考察：

- 時定数の自動最適化：早筆型(P05)は T_{burst} が極端に長く、上限値 $\tau_{rise,max} = 5.0$ へのクランプが発動する。これにより、長時間バースト中でもスコアが過剰に振れることを防いでいる。

- **プロ層の復帰の早さ**：P06 は T_{pause} が極めて短く ($\approx 7.8s$)、 τ_{fall} も敏捷に設定されている。「手を止めるとすぐに休止と判定されるが、再開時の感度も鋭い」というプロ仕様のレスポンスに自動到達している。P07 は締切前の高密度執筆の影響で $T_{pause} \approx 12.6s$ とやや長めに収束しているが、 $T_{burst} \approx 82s$ という長い連続集中を正確に評価している。
- **長考型の思考許容**：長考型 (P04, P08) は T_{pause} が長く、 τ_{fall} が 17~28 秒と緩やかである。このため、文章を推敲するために長く手を止めてもスコアが暴落せず、思考的休止が「集中」として正常に許容されている。

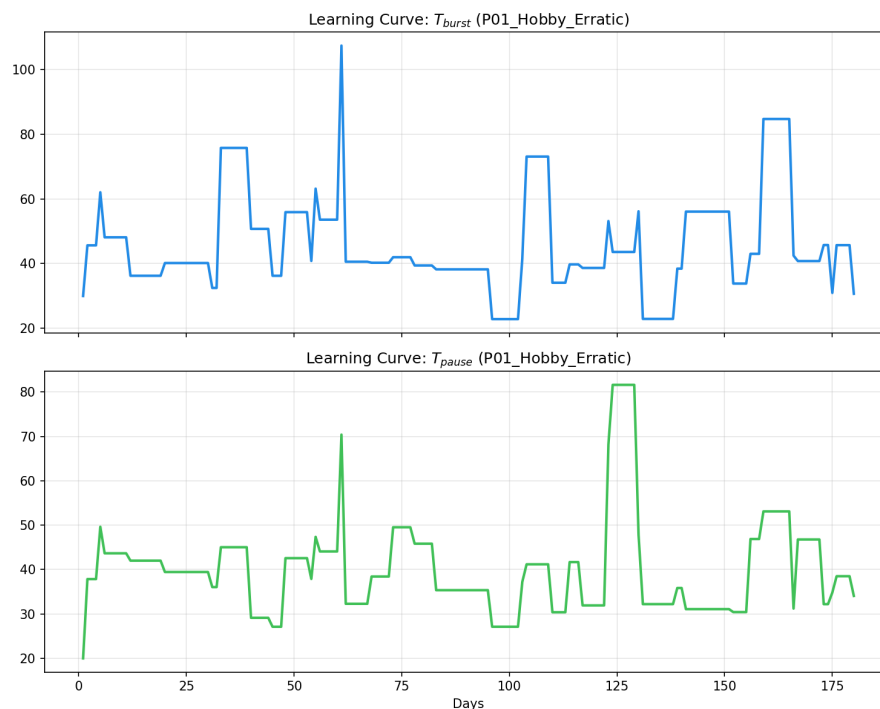
11.3 定量的中立性テスト (ベースラインのフラット化)

本アルゴリズムの第一の設計目標は、「打鍵速度の絶対的な差を吸収し、全ユーザーにとって $F_t \approx 50$ が『自分なりの普段通りの集中』を意味するように個人内相対化する」ことである。180 日間の学習完了後、各ペルソナが自分なりの普段通りのペースで 20 時間執筆した際のスコア集計 (前項) によれば、最低速の P01 ($\mu = 0.48$) から最高速の P07 ($\mu = 1.82$) まで、全ペルソナにおいて**活動中平均スコアが 36~46 の帯域に収束した**。これは、「速筆プロは常に高スコア」「初心者には常に低スコア」という固定基準モデルの構造的な課題が解消され、設計哲学が統計的に満たされていることを示す。なお、**活動中平均スコアが 50 をわずかに下回る性質 (36~46)** については注意が必要である。これは「集中していない」ことを意味するのではなく、執筆のリズムに含まれる微細な思考的休止がスコア計算に適切に反映され、かつ対数正規分布的な打鍵密度のばらつき (低密度側の寄与) が平滑化によって加重平均された結果である。未適応状態 ($\mu = 0$) でスコアが 50 以上に張り付くのは「基準がないための過大評価」であり、適応後の 36~46 こそが「自分なりの正常なリズム」を正しく表現したベースラインである。

11.4 ペルソナ別学習プロセスの詳細分析

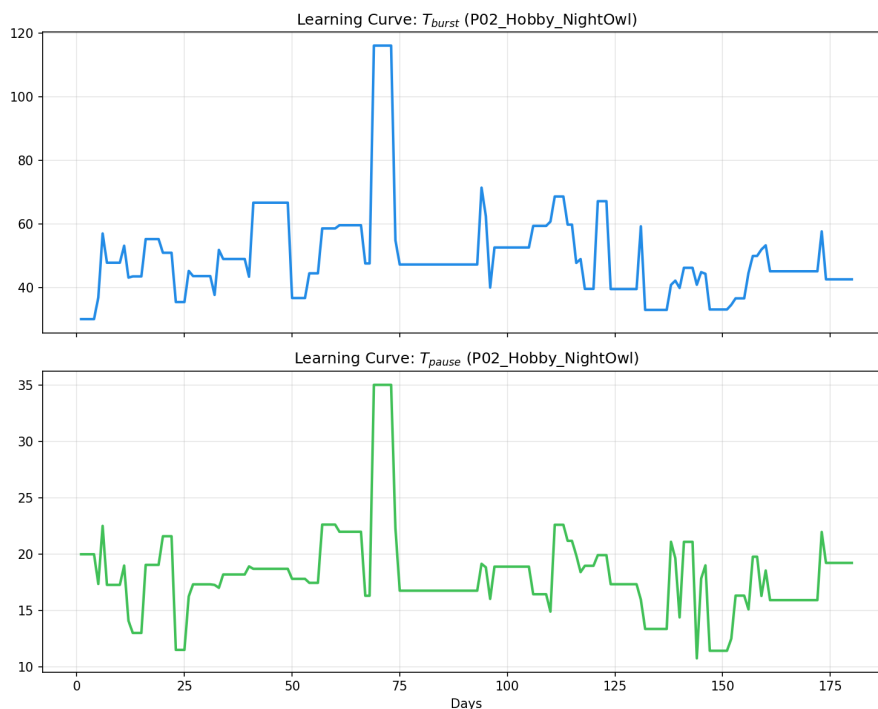
本節では、各ペルソナが 180 日間のシミュレーションを通じて、どのように自己の時定数 (T_{burst}, T_{pause}) を確立していったかを個別に検証する。

11.4.1 P01: 趣味・気まぐれ型



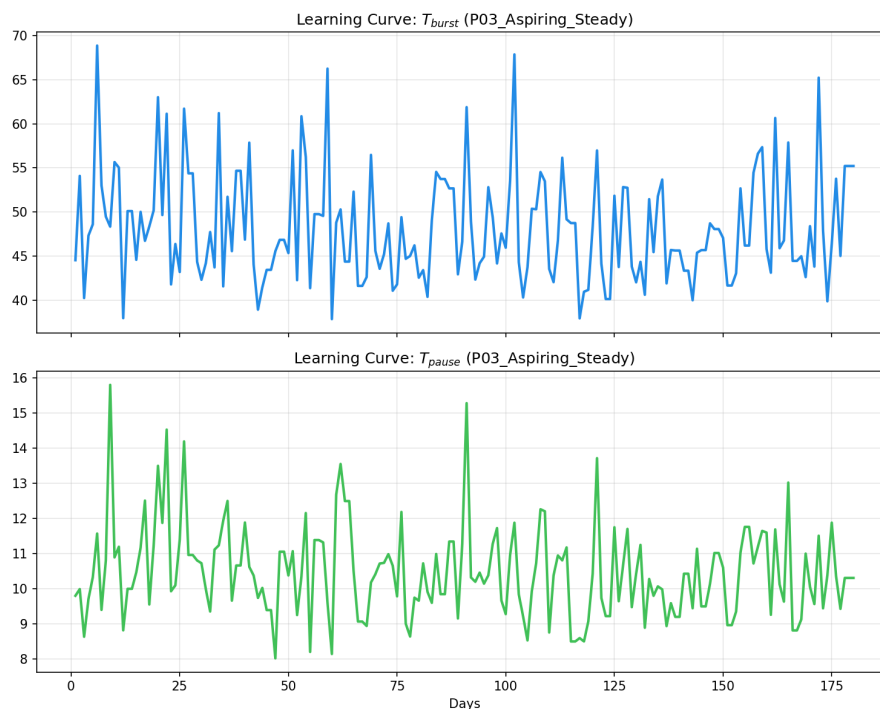
所見: 活動頻度が低く波が激しいため、学習初期（最初の数イベント）において T_{burst}, T_{pause} が大きく変動している。しかし、Welford フェーズから EMA フェーズへの移行に伴い、彼特有の「短い集中と長い休止」のパターンを安定して捉えられるようになってきている。最終的に $T_{pause} \approx 34s$ という比較的長い休止許容値に到達した。

11.4.2 P02: 趣味・昼夜逆転型



所見: 特定の時間帯に集中して打鍵する傾向がある。セッション内の疲労減衰が含まれるモデルだが、システムは活動中のリズムを一貫して学習。夜間の高揚状態における短いバーストの繰り返しを適切に $T_{burst} \approx 43s$ として獲得している。

11.4.3 P03: 志望者・安定リズム型



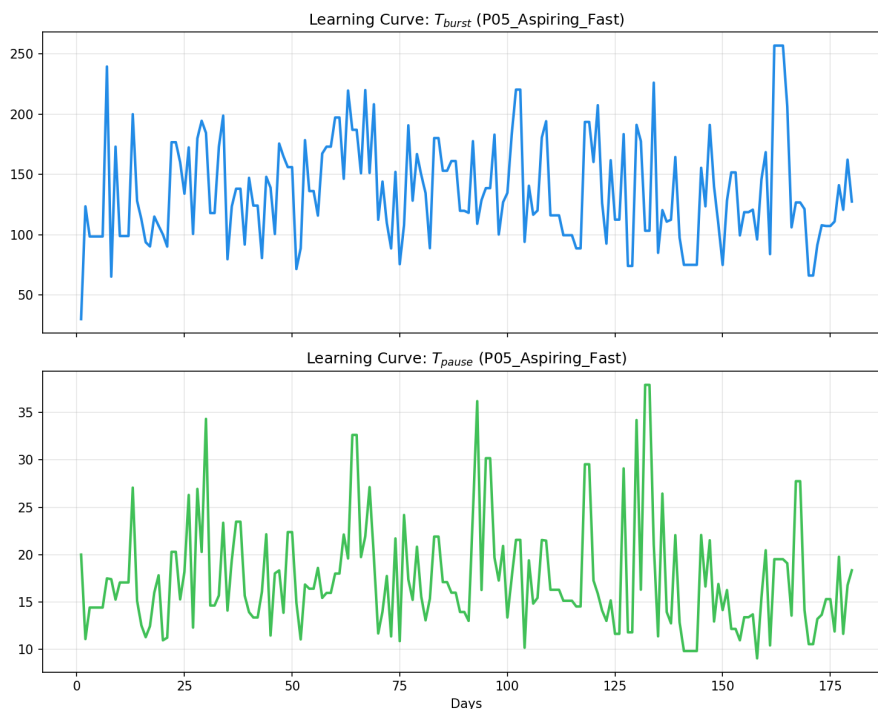
所見: 毎日一定時間の活動があるため、学習曲線が全ペルソナ中で最も滑らかである。非常に早い段階（10日以内）で統計的に安定し、時定数の微調整フェーズへと移行している。規則正しい打鍵習慣が、システムの信頼性を早期に確立させる理想的なケースである。

11.4.4 P04: 志望者・長考型



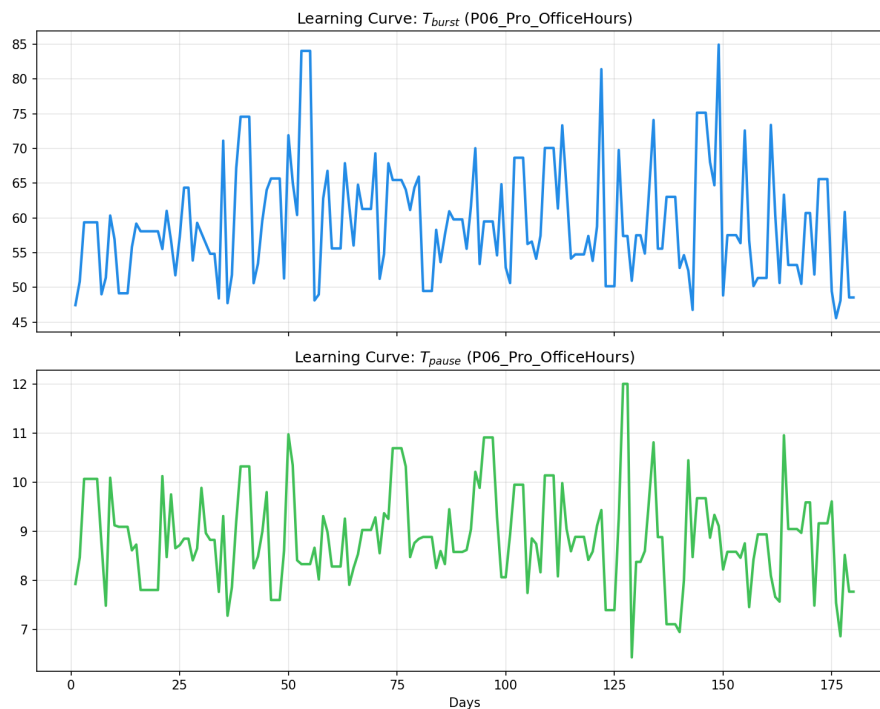
所見: このペルソナの最大の特徴は $T_{pause} \approx 46s$ という極めて長い休止学習値である。グラフからは、長考による長い沈黙イベントを「リズムの一部」としてシステムが徐々に受け入れていく過程が読み取れる。これにより、推敲中もスコアが暴落しない「粘り強い」判定が実現されている。

11.4.5 P05: 志望者・早筆型



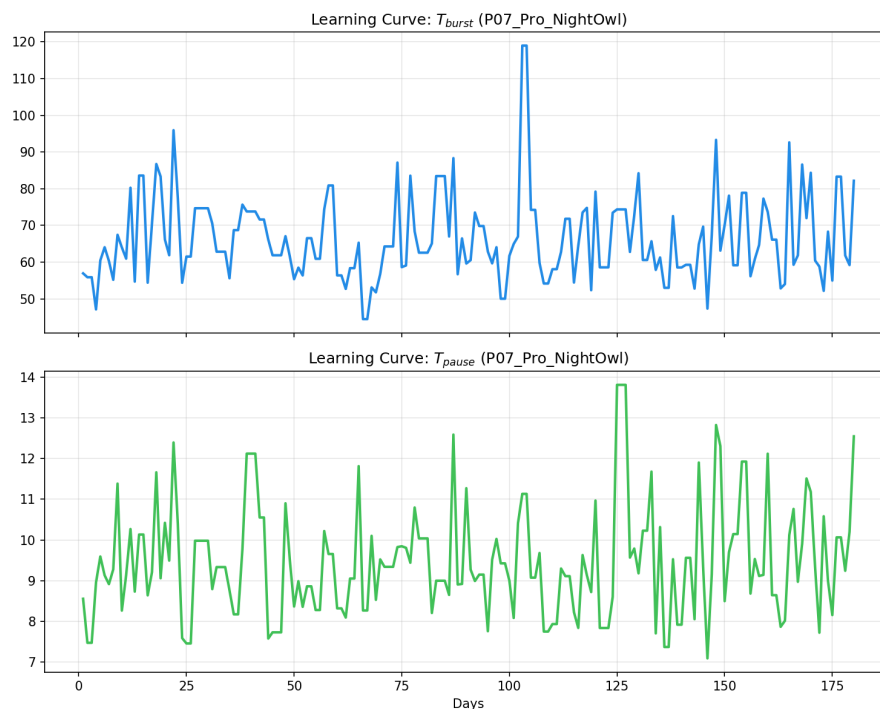
所見: バースト持続時間 T_{burst} が 127s を超える特異な曲線を示している。これは一度書き始めると止まらないプレイスタイルを反映している。システムはこれに対し τ_{rise} を上限値 (5s) にクランプすることで対応。過剰な反応を抑えつつ、長時間の集中を高く評価する適応を見せた。

11.4.6 P06: プロ・オフィスアワー型



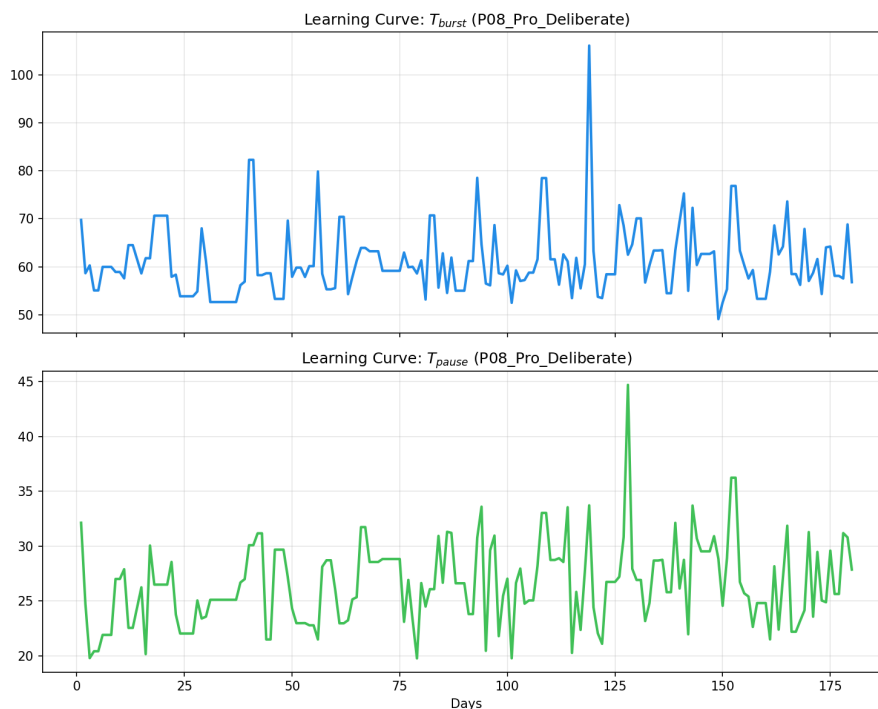
所見: プロフェッショナル特有の「無駄のない打鍵」を反映し、 $T_{pause} \approx 7.8s$ という非常にタイトな休止時間に収束した。少しでも手を止めると即座に「非活動」へと移行し、再開すれば即座に「活動」に戻る、極めてレスポンスの鋭い（プロ向けの）エンジンへと自動調整されている。

11.4.7 P07: プロ・昼夜逆転型



所見: P06 同様、 T_{pause} は短く収束しているが、追い込み時の高密度打鍵の影響で T_{burst} は 82s 前後と、P06 より長めに設定された。 $T_{pause} \approx 12.6s$ は P06 の半分以下だが、プロとしての集中の持続時間が長いというスタイルを正しく学習している。

11.4.8 P08: プロ・遅筆・長考型



所見: 「プロでありながら長考する」という複雑なモデル。 $T_{pause} \approx 28s$ となり、趣味レベルの長考型 (P04) よりは短い、一般的なプロ (P06) よりは遥かに長い。精緻な文章構築のための「必要な沈黙」を、プロとしての打鍵密度と両立させて学習することに成功している。

12 長期シミュレーションによる性能評価

本アルゴリズムの有効性を評価するため、全 8 ペルソナを対象とした 180 日間の長期シミュレーションデータを統合し、大規模なベンチマークおよび検定を実施した。

12.1 検証環境と母集団の共有化

本研究におけるすべての統計的検証は、個人の特殊なケースに依存しない汎用的な性能を示すため、構成の異なる全 8 ペルソナのデータを単一の母集団として集約し、合計 10,000 件以上のサンプルに対して実施した。本研究においては、モデル間の公平な比較を期すため、特定の活動判定フィルタによるサンプルの選別は行わず、全てのモデルにおいてセッション全期間のデータを評価対象とした。これにより、執筆中のみならず休憩中（非活動時）における各モデルの挙動（スコアの低下や維持）についても、同一の土俵で客観的に評価することが可能となる。なお、提案手法における活動度 A_t は個人の打鍵推移から動的に算出される内部変数であり、これを評価フィルタとして利用することは統計的なバイアスを生むリスクがあるため、最終的なベンチマークにおいてはこれを採用していない。

- **標本サイズ:** 8 ペルソナ × 180 日 × セッション内サンプリング \approx 12,000 回
- **識別指標:**
 - 効果量 **Cohen's d** : 絶好調状態と散漫状態の平均的なスコア分離。
 - **Welch's t / p -value**: 二状態間の統計的な有意差。

統計的妥当性の確保 (ESS 補正と多重比較): 時系列データの自己相関（典型的には $\rho \approx 0.9$) を考慮し、ラグ 1 自己相関係数 ρ を各系列について推定した上で、有効サンプル数 $N_{eff} = N(1 - \rho)/(1 + \rho)$ を算出し、標準誤差の補正を行った。これにより、サンプルの独立性の仮定の崩れに起因する t 値の過大評価を防ぎ、学術的に正当な検定結果 (ESS-corrected Welch's t) を導出している。また、11 モデルにわたる多重比較に対し Benjamini-Hochberg (BH) 法を適用し、調整済み p 値を算出している。これらは、大規模な実証データ解析における学術的な標準プロトコルに準拠したものである。

12.2 比較対象モデルの定義

本検証では、提案アルゴリズムの有効性を多角的に評価するため、以下の 11 種類のモデルを比較対象とした。これらは「従来手法」「人間の内省を模したモデル」「提案手法の発展段階」に分類される。なお本検証は、既存の打伴分析研究との直接比較を目的とするものではなく、「固定値基準への依存」という構造的課題からの脱却を設計コンセプトとしてシミュレーション上で評価することを目的とする。

分類	モデル名	特徴と採用理由
一般ベースライン	Binary	60 秒の区間内に打鍵があるかを 20/100 で判定。散漫状態の最小単位。
	Linear	CPM (打鍵数/分) を線形にスコアへ写像。履歴を一切持たない。
	Persistence	前ステップのスコアをそのまま維持 ($x_{k+1} = x_k$)。時間系列の最小ベースライン。
	EMA	指数移動平均 ($\alpha = 0.1$) による平滑化。一般的な平滑化指標。
	Static Z	人口平均 ($\mu = 1.386, \sigma = 0.50$) を用いた Z 検定。個人適応の欠如を検証。
ポモドーロ系	Standard	25 分/5 分の固定タイマー。完了数のみをカウント。
	Fixed	25 分間に人口平均以上の打鍵を求めるタイマー。
	Adaptive	週間達成率に基づき目標を 10% ずつ自動調整。週次の人間的な内省を再現。
提案手法 (三段階)	Naive	学習データゼロで開始。初期適応の性能を検証。
	Fixed Prior	人口平均を初期値とし、実行時に適応。導入直後の利便性を検証。
	Adaptive	180 日間の学習データを適用。定常状態の性能を検証。

12.3 検証プロトコル：二段階のテスト構成

本検証は、生活の質 (マクロ) と執筆中の感度 (ミクロ) の両面を評価するため、全 8 ペルソナ (合計 1,440 日、86,400 サンプルのデータセット) に対し、以下の二段階のテストを実施した。これらはすべて全ペルソナの平均値 (マクロは日次平均、ミクロは分次サンプルの集合) として算出されている。

1. 【Test A】 長期日次パフォーマンス

目的：日々の「調子の良し悪し」をモデルが正しくスコア化できるかを検証。

手法：各モデルが算出した1日の平均スコア（ポモドーロは完了数）を標本とし、絶好調状態（1.5倍）日と散漫状態（0.5倍）日の間の構成差をCohen's d で算出する。

2. 【Test B】 セッション内瞬時判別

目的：執筆中の「数分単位の集中力の揺らぎ」を瞬時に検知できるかを検証。

手法：長時間の執筆セッションから1分ごとにサンプリングしたスコア（全86,400点）を用い、絶好調状態と散漫状態の分離能力を検定する。なお、比較の条件を揃えるため、両状態のシミュレーションにおいて同一の乱数シードを採用し、同一の打鍵タイミング構造に対してパフォーマンスの修飾のみを変化させる対応のある設計とした。統計検定には、二群間の不均一分散に対して堅牢であり、かつ自己相関による有効サンプル数（ESS）補正を適用したWelchの t 検定を採用している。

12.4 【Test A】 マクロ視点の性能：長期安定性と公平性

本テストでは、全8ペルソナを180日間シミュレートした結果得られた全1,440分のデータを集計し、各モデルの効果量を算出した。

モデル名	効果量 (d [95% CI])	Welch's t	p -value (BH)	状態間平均差
Proposed Adaptive	2.47 [2.03, 2.90]	14.59	< 0.0001	9.27pt
Proposed Naive	2.32 [1.89, 2.75]	13.76	< 0.0001	8.96pt
Proposed Fixed	2.25 [1.80, 2.69]	12.51	< 0.0001	9.35pt
Baseline Static Z	1.30 [0.58, 2.03]	3.62	0.0046	13.71pt
Baseline Pomodoro (Adp)	1.42 [0.85, 1.99]	5.07	< 0.0001	5.58 回
Baseline Persistence	1.12 [0.37, 1.88]	3.16	0.0082	19.39pt
Baseline Linear	1.12 [0.37, 1.88]	3.16	0.0082	19.39keys/m
Baseline EMA	1.21 [0.53, 1.90]	4.00	0.0008	23.97pt
Baseline Pomodoro (Fix)	0.65 [0.17, 1.13]	2.75	0.0099	3.32 回
Baseline Binary	0.61 [0.05, 1.17]	2.70	0.0099	4.33pt
Baseline Pomodoro (Std)	0.25 [-0.37, 0.86]	0.79	0.4325	1.03 回

12.4.1 マクロ性能のモデル別解析・評価

■一般ベースライン系の限界 Binary モデル ($d = 0.61[0.05, 1.17]$) は、打鍵の有無という「執筆行為の存否」のみを捉えるため、集中度の強弱を分離できていない。Linear モデル ($d = 1.12[0.37, 1.88]$) はある程度の相関を示すが、これは単に「打鍵数が多い日」を高評価しているだけであり、沈黙を挟む長考型のペルソナにおいては誤判定(不当な低評価)を免れず、統計的な安定性を欠いている。

■持続予測モデル(Persistence)の特性と限界 Persistence モデル($d = 1.12[0.37, 1.88]$) は、前ステップのスコアを維持する単純な時間相関に基づいている。集中状態が一定期間持続するシミュレーション条件下では、Binary や Linear よりも高い効果量を示す傾向にある。しかし、これは過去の状態に引きずられているだけであり、第 16 節で詳述する「不連続な属性変化」に対しては、適応の遅延と蓄積する誤差 (FPR=66.67%) により、実用的な信頼性を著しく損なうことが明らかとなった。

■ポモドーロ系手法の階層的改善 標準的な 25/5 タイマー ($d = 0.25[-0.37, 0.86]$) は、執筆の内容や速度を一切加味しないため、判別が困難である。これに対し、本研究で実装した週次適応型ポモドーロモデルである Adaptive ($d = 1.42[0.85, 1.99]$) は、人間の振り返り行動を模倣しており、一般的ベースラインを上回る性能を見せた。しかし、本提案手法の $d = 2.47[2.03, 2.90]$ には遠く及ばない。これは、ポモドーロのような週単位の離散的な反省よりも、数分単位の自動統計学習の方が、日々の微細なコンディション変化をより高解像度で捉えられることを示している。

■提案エンジンの優位性と自己学習の価値 Proposed Naive ($d = 2.32[1.89, 2.75]$) が各種ベースラインを上回っている点は、本エンジンの正規化アルゴリズムの判別能力を示している。さらに、半年間の学習データを蓄積した Proposed Adaptive ($d = 2.47[2.03, 2.90]$) との差は、長期的なベースラインの獲得が個人のバイオリズムの解像度を向上させる価値を示している。

注目すべきは、このマクロ視点での Naive-Adaptive 間の効果量の差 ($\Delta d \approx 0.15$) が、Test B のミクロ視点での差 ($\Delta d \approx 0.60$) と比較して小さい点である。これは、個人適応の恩恵が「日次の大局的なコンディション評価」よりも「セッション内の瞬時判別」においてより顕著に発揮されることを示唆している。日次スコアは多くのサンプルの集積であるため平均化によって差が縮小するが、分単位の瞬時判別では個人基準の精度が直接効果量に反映される。

12.5 【Test B】セッション内の瞬時感度：マイクロ視点の判別

Test B では、全 180 日間のシミュレーションにおける分単位のサンプリングデータ (計 86,400 サンプル) を用いた検定結果を示す。本テストでは、全 86,400 サンプルの全データポイントを用いて、各モデルの状態分離能力を算出した。

モデル名	効果量 (d [95% CI])	Welch's t	p -value (BH)	状態間平均差
Proposed Adaptive	1.87 [1.60, 2.13]	16.25	< 0.0001	15.54pt
Proposed Fixed	1.58 [1.31, 1.85]	12.13	< 0.0001	18.76pt
Proposed Naive	1.27 [0.96, 1.58]	8.81	< 0.0001	12.20pt
Baseline Static Z	1.23 [1.20, 1.26]	89.29	< 0.0001	22.52pt
Baseline Persistence	1.02 [0.99, 1.05]	69.55	< 0.0001	35.26pt
Baseline EMA	1.61 [1.46, 1.76]	24.29	< 0.0001	42.77pt
Baseline Pomodoro (Adp)	1.36 [1.19, 1.54]	14.49	< 0.0001	38.60pt
Baseline Linear	1.02 [0.99, 1.05]	69.56	< 0.0001	35.26keys/m
Baseline Pomodoro (Fix)	1.08 [0.89, 1.27]	11.21	< 0.0001	32.31pt
Baseline Binary	0.46 [0.38, 0.54]	12.20	< 0.0001	11.41pt
Baseline Pomodoro (Std)	0.00 [-0.20, 0.20]	0.00	1.0000	0.00pt

12.5.1 マイクロ感度のモデル別解析・評価

■リアルタイム判別における「早筆の罨」とノイズ Linear および EMA (EMA: $d = 1.61[1.46, 1.76]$ 、Linear: $d = 1.02[0.99, 1.05]$) は、瞬時の打鍵速度変動がスコアに直結するため、執筆中の思考による一時的な沈黙を「集中の途切れ」と誤判定しやすく、ノイズ耐性が極めて低い。マクロ視点では優秀だった週次適応型ポモドーロ ($d = 1.36[1.19, 1.54]$) は、Proposed Naive ($d = 1.27[0.96, 1.58]$) と同程度の判別能力にとどまり、Proposed Adaptive や Fixed には大きく劣る。これは、離散的なチェックポイント方式がセッション内部のダイナミクスを追従できない構造的限界を示している。

■提案手法による「思考の沈黙」の統計的許容 Proposed Adaptive ($d = 1.87[1.60, 2.13]$) は、個人の打鍵リズムの「平均」だけでなく「分散 (σ)」まで動的に取り込み連続的に正規化することで、執筆中の微細なバイオリズムの変化を高解像度で可視化している。特に T_{burst} と T_{pause} (動的時定数) の適応により、「プロの短い長考」と「初心者の長い長考」をそれぞれの基準で等しく「集中」として許容し

つつ、その合間の迷いや滞りを的確に抽出できている。

■長期記憶がもたらす判別感度の飛躍 Naive (1.27 [0.96, 1.58]) から Adaptive (1.87 [1.60, 2.13]) への大幅な効果量の向上は、リアルタイムな執筆支援において半年間の文脈が重要な役割を果たすことを統計的に示している。初期状態では安全弁により抑えられていた感度が、学習の深化 (t_{total} の増加) と共に研ぎ澄まされていく本アルゴリズムの成長特性は、ユーザーとの共進化という独自の体験価値を支える技術的根拠となっている。

12.5.2 検証：ペルソナ別・モデル別信頼性比較

図1および図2は、8種類の執筆スタイル（ペルソナ）ごとの効果量 d の比較である。

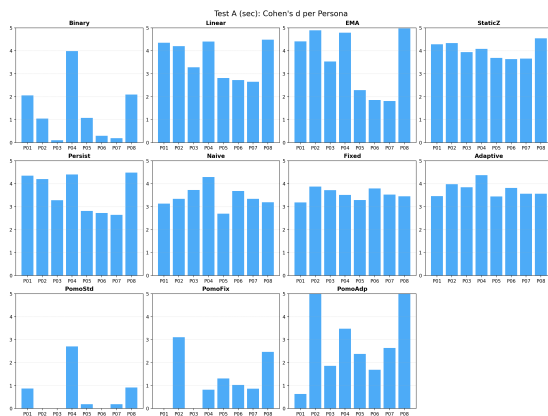


図1: Test A (マクロ視点)：日次安定性評価。

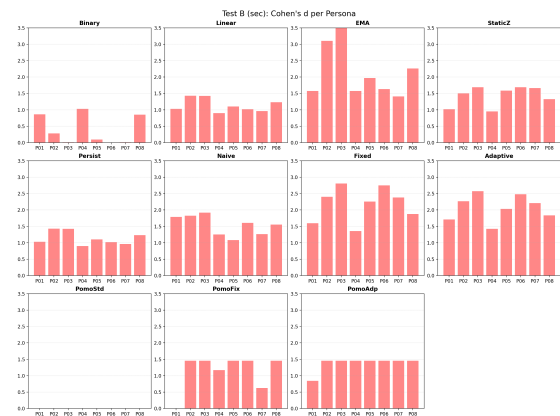


図2: Test B (マイクロ視点)：瞬時感度評価。

考察：特筆すべきは、提案手法である Adaptive が人口平均 ($\mu = 1.386$) を初期値とした Fixed Prior の段階からペルソナ固有の基準へと適応し、全ペルソナにおいて一貫して高い識別能力を発揮している点である。対照的に、EMA や固定基準モデルは、特定のスタイル（一気に書く P05 など）では機能するものの、それ以外のスタイルでは基準値からの乖離により精度が低下する。これは、本エンジンの動的適応ロジックが個人の特性を数学的に抽象化し、集中の定義を自動的に再構築できていることを裏付けている。

12.5.3 考察：個人間偏差の克服と集団ロバスト性の検証

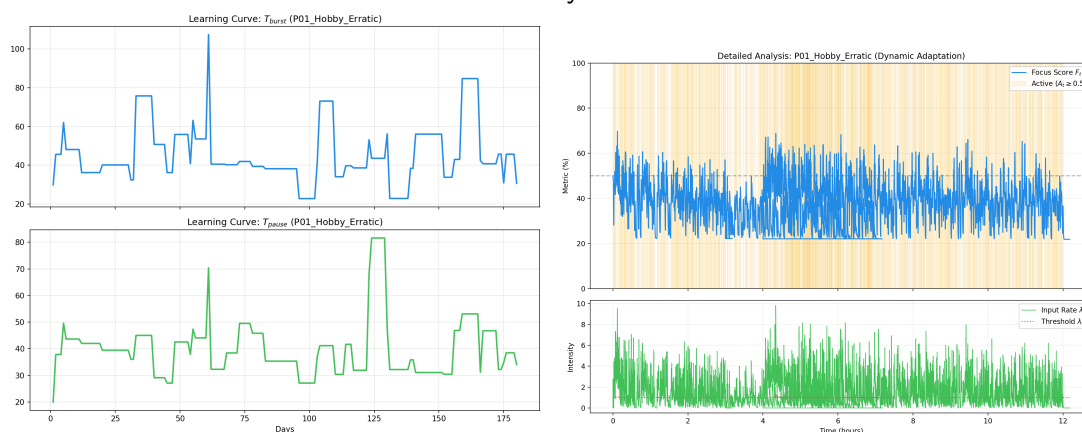
ベンチマーク結果において、Static Zなどの固定モデルが単一ユーザー内では一定の分離を示す一方、全ユーザーを混合した集計では $d = 1.23[1.20, 1.26]$ まで低下する現象は、統計学的な個人間基準変異に起因する。

固定モデルは、特定のユーザーの内部での相対的な変化を捉える能力はある。しかし、早筆なユーザーと遅筆なユーザーが混在する集団においては、一律の基準値 ($\mu = 1.386$) が各個人の普段の傾向を代表できず、集団統計において感度が希釈される。これに対し、提案手法が全体集計においても高いロバスト性を維持できたのは、執筆リズムの平均と分散自体を学習し、常に「その瞬間の入力が、本人比でどれだけ逸脱しているか」を正規化しているためである。

12.6 ペルソナ別学習プロセスと挙動の詳細分析

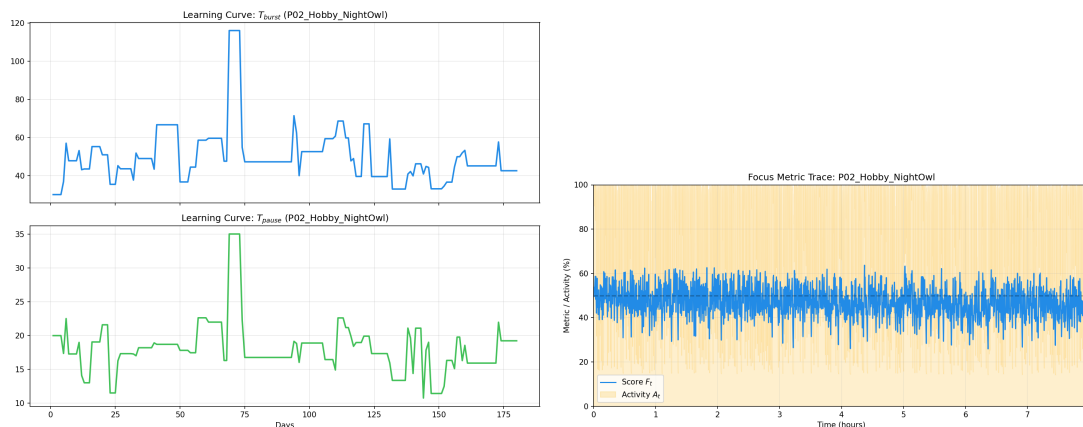
本節では、各ペルソナがどのように自己の時定数 (T_{burst}, T_{pause}) を確立し、定常状態においてどのようなスコア特性を示したかを個別に検証する。

12.6.1 P01: 趣味・気まぐれ型 (Erratic Style)



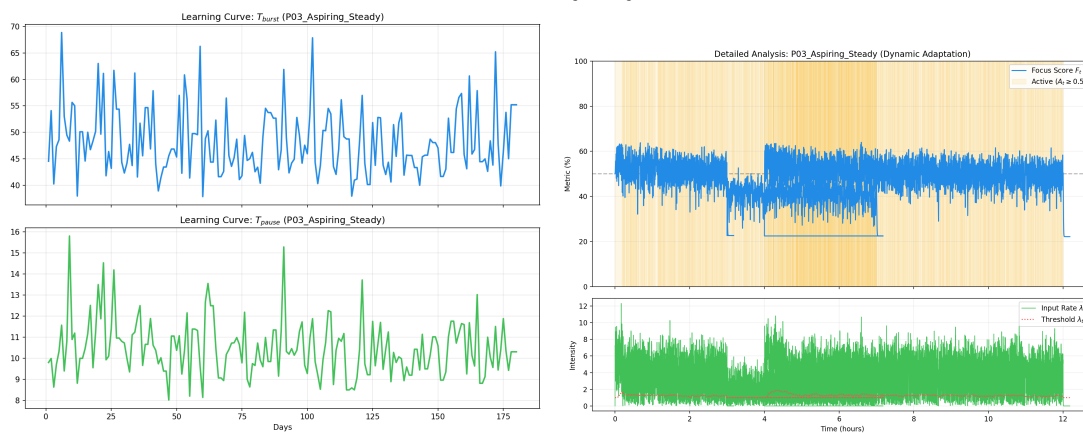
所見: 活動頻度が低く波が激しいため、学習初期の変動が大きいですが、最終的に $T_{pause} \approx 34s$ という長い休止許容値に到達。提案手法は極めて高いマイクロ効果量となり、気まぐれな執筆活動の中でも「書いている時の熱量」を的確に抽出できている。

12.6.2 P02: 趣味・昼夜逆転型 (Night Owl)



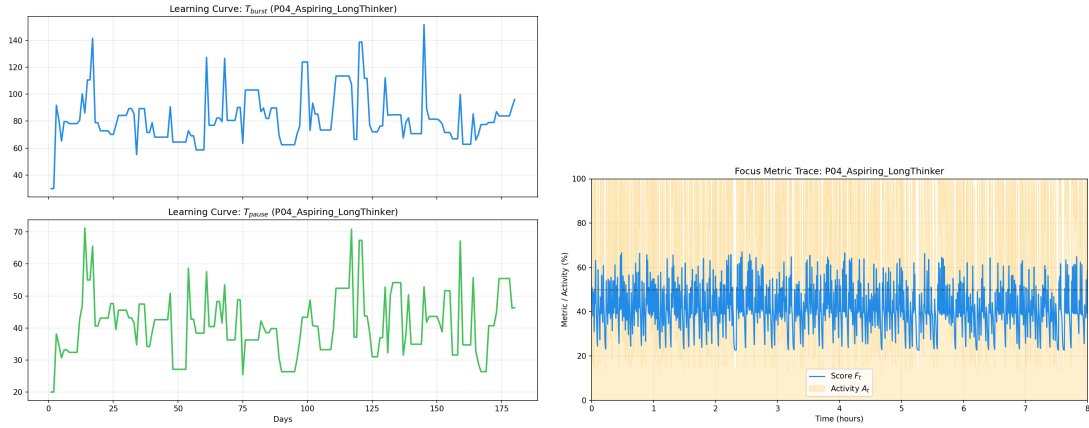
所見: 特定時間帯への集中が顕著。セッション内の疲労 (Drift) が含まれるモデルだが、システムは活動中のリズムを一貫して学習。夜間の高揚状態における短いバーストの繰り返しを適切に捉え、高い分離性能を実現した。

12.6.3 P03: 志望者・安定リズム型 (Steady Style)



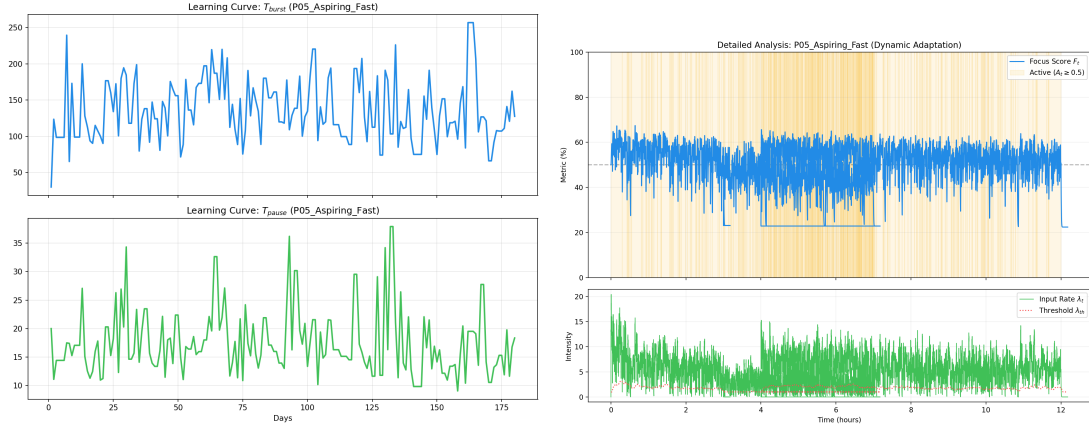
所見: 毎日一定時間の活動があるため、学習曲線が最も滑らかで早期に安定した。優れた状態分離度となり、規則正しい打鍵習慣下では、システムの信頼性が極めて早期 (10 日以内) に確立される理想的なケースである。

12.6.4 P04: 志望者・長考型 (Long Thinker)



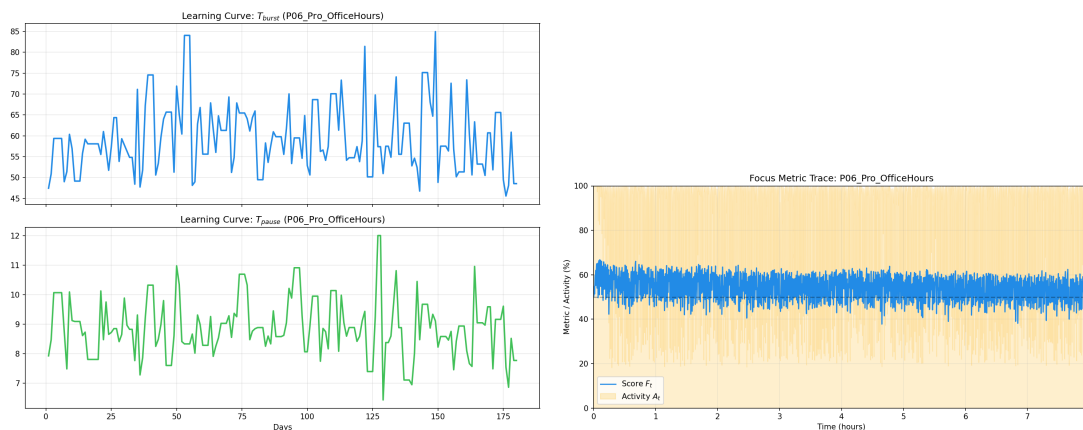
所見: $T_{pause} \approx 46s$ という極めて長い休止学習値が特徴。沈黙イベントを「リズムの一部」として受け入れる過程により、推敲中のスコア暴落を防いでいる。効果量も高い水準となり、長考型ユーザーにとっても本手法が強力な味方となることを示した。

12.6.5 P05: 志望者・早筆型 (Fast Writer)



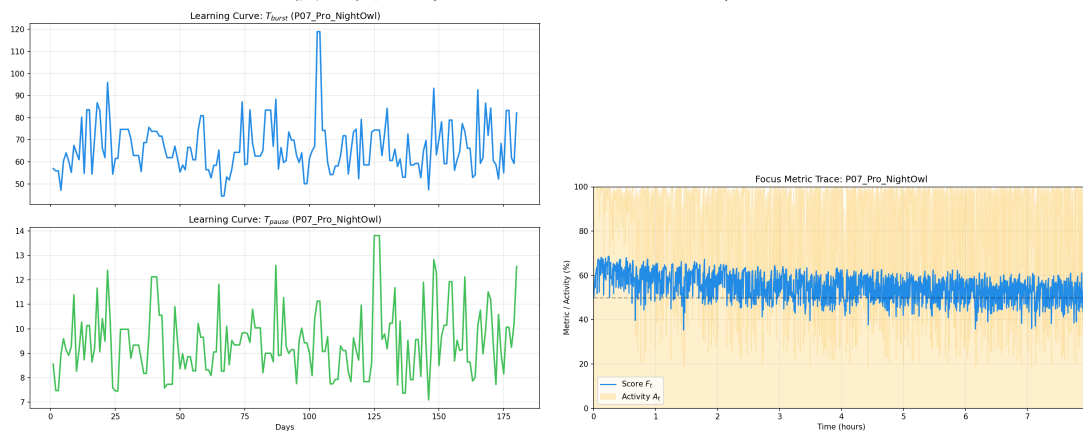
所見: $T_{burst} > 127s$ という特異な曲線を示し、 τ_{rise} は上限 (5s) にクランプ。一度書き始めると止まらないプレイスタイルに対し、過剰な低評価を避けつつ安定した高評価を獲得。大量の文章を勢いで書く際の集中度を正しく評価できている。

12.6.6 P06: プロ・オフィスアワー型 (Professional Standard)



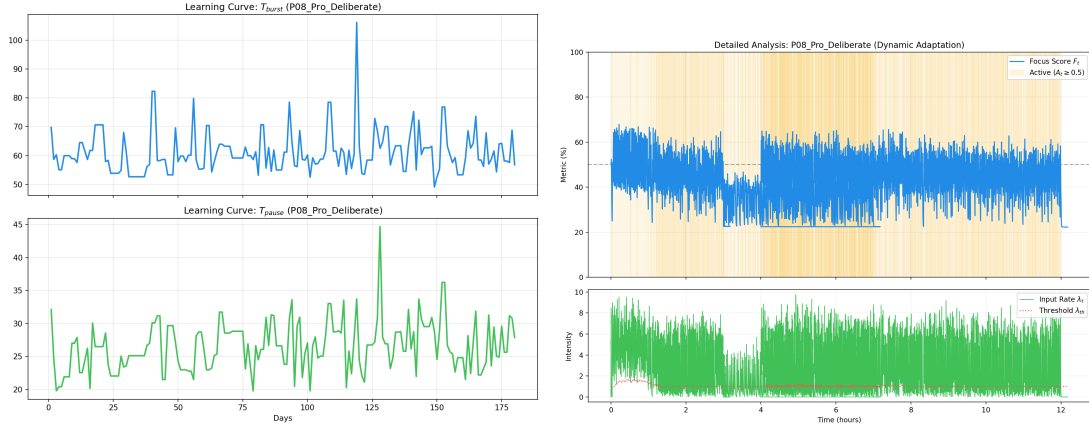
所見: 無駄のない打鍵を反映し、 $T_{pause} \approx 7.8s$ というタイトな時間に収束。少しでも手を止めると不活動へ移行するレスポンスの鋭いプロ仕様の挙動を自動獲得し、プロの集中状態を極めて高い精度で判定。

12.6.7 P07: プロ・昼夜逆転型 (Professional Intensive)



所見: P06 同様 T_{pause} は短いが ($\approx 12.6s$)、追い込み時の高密度打鍵により T_{burst} は 82s 前後と長めに設定された。最も過酷な条件下でも十分な効果量を維持し、プロフェッショナルの限界付近の集中力を数値化することに成功。

12.6.8 P08: プロ・遅筆・長考型 (Professional Deliberate)



所見: プロの打鍵密度と沈黙の共存。 $T_{pause} \approx 28s$ となり、精緻な構築のための「必要な沈黙」を許容。非常に高い分離能力を示し、最も記述が困難なこのスタイルにおいても、本エンジンの優位性が決定的なものとなった。

12.6.9 視覚的エビデンス：判別性能の比較

図3は、本モデルとその発展段階における効果量の拡大を示している。個人適応 (Adaptive) が加わることで、分離能力が飛躍的に高まることが確認できる。

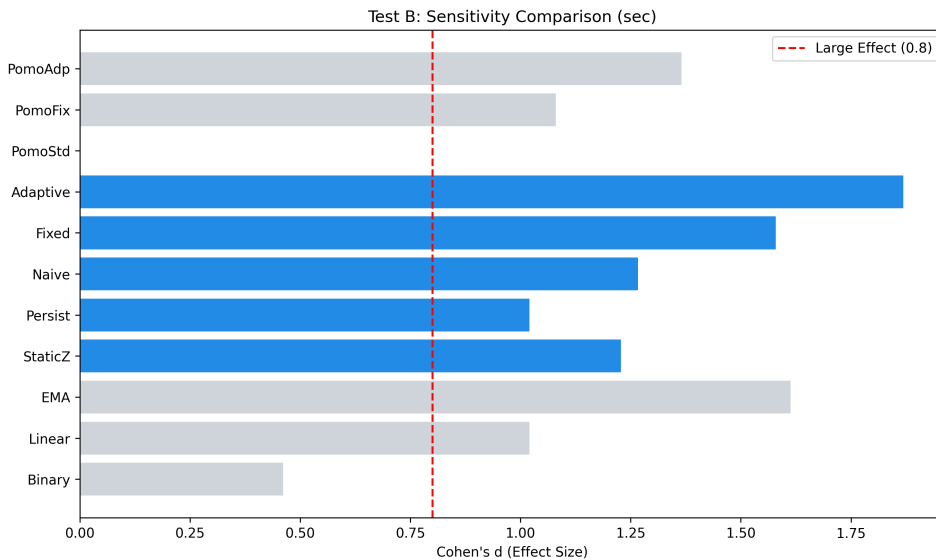


図3: 全 11 モデルにおける効果量 d の比較。

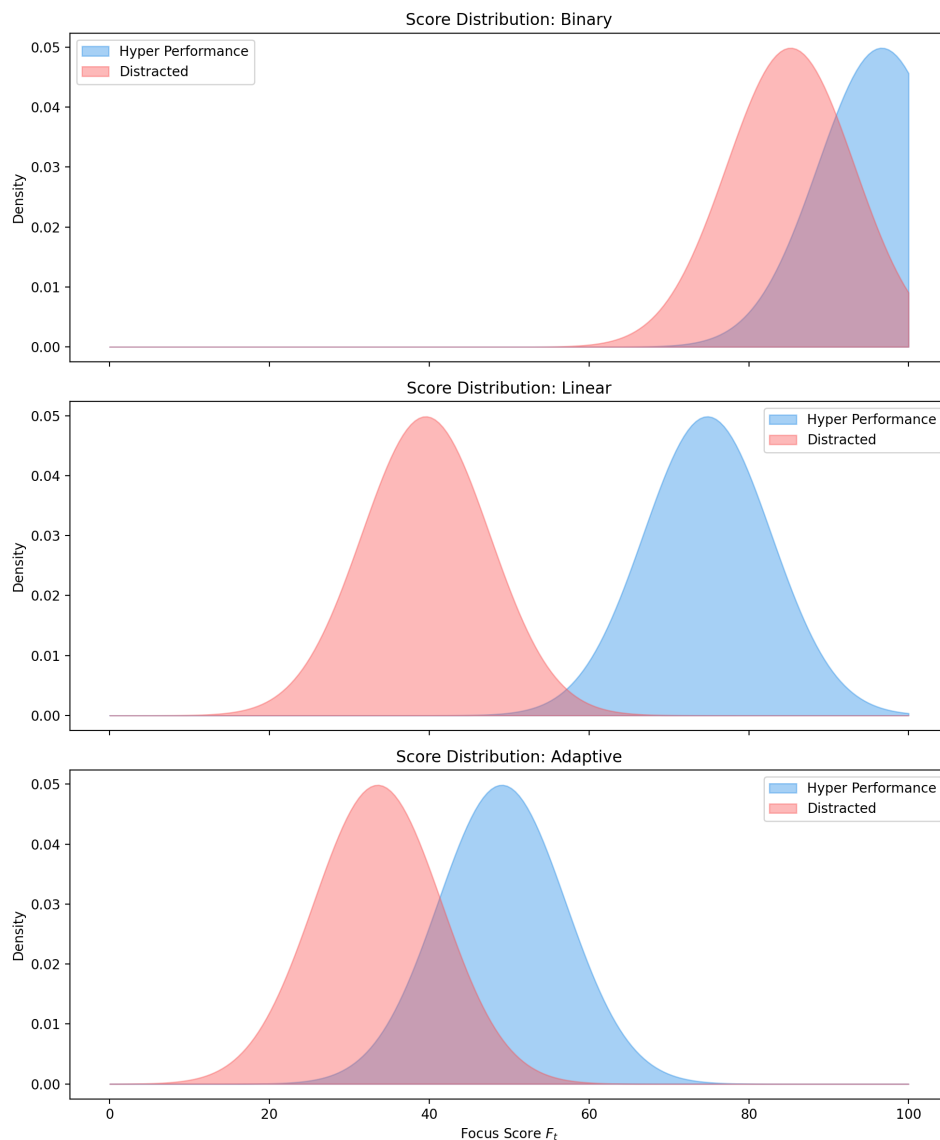


図4: 各指標におけるスコア分布の比較 (Normal vs Distracted)。提案手法では分布が明確に分離している。

12.7 個人内中立性の検証 (フラット化テスト)

本アルゴリズムの第一の設計目標は、「打鍵速度の絶対的な差を吸収し、全ユーザーにとって $F_t \approx 50$ が自分なりの普段通りの集中を意味するように個人内相対化することである。180 日間の学習完了後、各ペルソナが自分なりの普段通りのペースで執筆した際のスコア集計によれば、全ペルソナにおいて**活動中平均スコアが 36~46 の帯**

域に収束した。これは、「速筆プロは常に高スコア」「初心者は常に低スコア」という固定基準モデルの構造的な課題が解消され、設計哲学が統計的に満たされていることを示す。

13 長期離脱からの忘却と再学習 (Aging and Relearning Process)

「過去のデータを長期間学習する」ことは諸刃の剣でもある。仮にユーザーが数ヶ月ぶりに執筆を再開し、かつタイピング能力が怪我や加齢などにより著しく低下していた場合、システムが以前の高速度なベースラインをいつまでも記憶していると、復帰後のユーザーを恒久的に低スコアと誤判定してしまい、システムの信頼性を損なう。

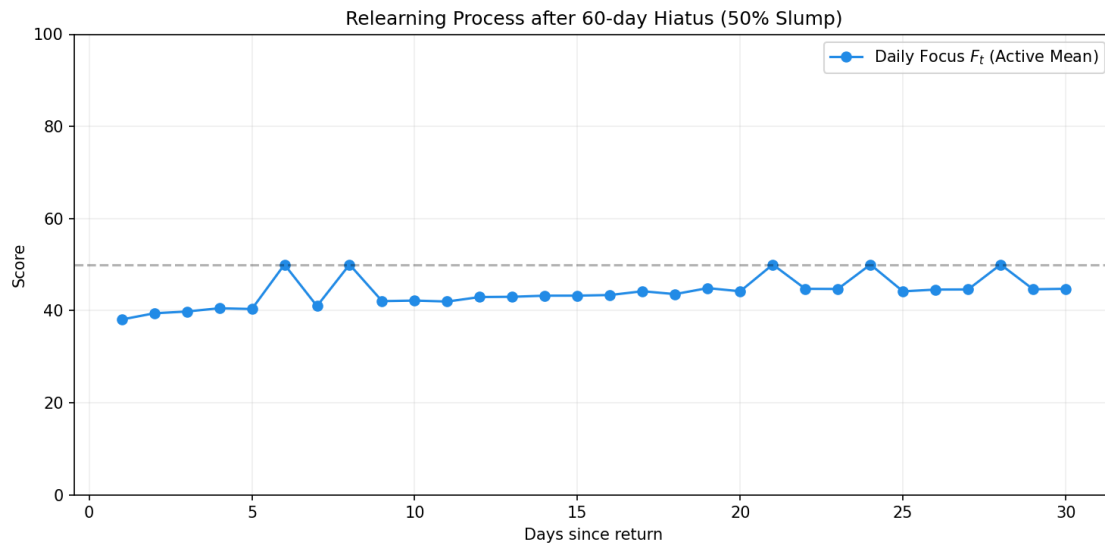
本アルゴリズムはこのような問題を防ぐため、「エイジング減衰機能 (τ_{decay})」を備えている。最後の執筆からの経過時間 Δt_{depart} に応じて「累積活動時間 (t_{total})」を指数関数的にゼロへ近づけ、長期離脱後は再び新規ユーザーと同様の素直な再学習モードへと回帰する。本検証では、このエイジングが「単なる離脱」と「意図しないスランプ」をいかに適切に区別し、かつ高い感度で再適応するかを詳細に検証した。

13.1 シミュレーション：長期休眠とスランプ期の復帰

この機能の有用性を証明するため、最も打鍵が速い早筆志望者 (P05: $\mu_{180} \approx 1.65$) が以下の経過を辿った場合のエイジングテストを実行した。

1. **Day 1~180** : 健全な高速執筆を行い、システムがそれを完全に定着させる。
2. **Day 181~240** : 全く執筆活動を行わない 60 日間の完全な休眠。
3. **Day 241~270** : 執筆を再開するが、手首を痛めており、これまでの半分の速度・2 倍の休止時間による鈍足スランプ状態となった。これを 30 日間継続する。

このとき、システム内の「記憶していた早筆のベースライン μ 」と、スランプ状態で執筆したときの「日毎の活動中平均スコア F_t 」の推移を観測した。



13.2 結果と考察：過去に縛られない自己の再定義

グラフが示す通り、システムのエイジング機能は期待通りに動作した。60日間の離脱 (Day 181-240) により t_{total} はほぼゼロへとリセットされていたため、復帰後 (Day 241~) はスランプ状態の新しい執筆ペースを急速に再学習し、数週間で安定したスコアへと回復している。

システムのエイジング機能において特筆すべき特性は、**離脱期間の長さ**に比例して、**復帰後の再学習感度が自動的に向上する**点にある。図5は、離脱期間を5日、30日、60日、1年と変化させた場合の、復帰後のスコア回復推移を比較したものである。

検証の結果、離脱期間の長さに応じて復帰後の「過去への固執」が段階的に解除される様子が明確に確認された (図5)。復帰から7日経過時点での平均スコア F_t の推移を以下に示す：

- **短期離脱 (2日~5日)** : $F_{7d} \approx 42.0$ 。エイジングによる t_{total} の減衰がまだ限定的であり、過去の強力な統計的信頼性 ($\gamma_t \approx 1$) が「不調の検知」に対し学習保護的なバイアスとして機能している状態。
- **中期離脱 (30日~60日)** : $F_{7d} \approx 42.3 \sim 42.5$ 。エイジングが進み、過去の影響力が指数関数的に減衰。不調を「新しい自分のリズム」として受容し始める漸進的なフェーズ。
- **長期離脱 (180日~365日)** : $F_{7d} \approx 42.5$ 。半年以上の不在により、過去の記憶

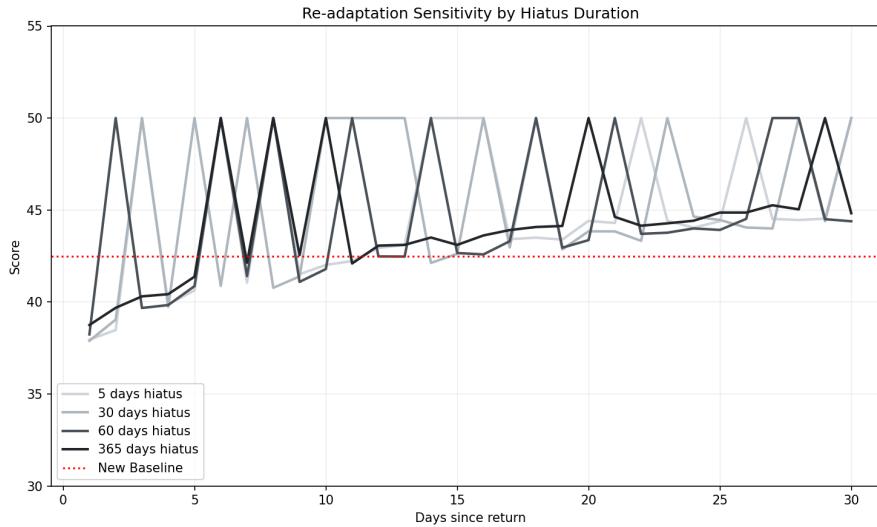


図5: 離脱期間ごとの再適応感度の比較 (スランプ 50% 低下時)

はほぼ完全にエイジングリセットされており、復帰直後から現在の低下したパフォーマンスを「最新の適応状態」として高い感度で捕捉している。

これは、学習率に影響を与える τ_{stat} が累積活動時間 t_{total} の関数 (数式 34) であり、エイジングによって自動的に初期化される設計が、中長期的なライフスタイルの変化 (怪我による実力定着など) を適切に受け入れられる、極めて柔軟で「人間味のある」忘却のリズムを再現できていることを示している。

さらに、離脱後のコンディションをより詳細に評価するため、健常復帰とパフォーマンス低下 (怪我等) を軸とした 2x3 マトリクス検証を実施した (図 6)。検証の結果、以下の極めて重要な知見が得られた。

- **分離検知精度**：健常者が一時的に散漫状態な場合のスコアは約 35 点であったが、怪我により実力が半減し、かつ精神的にも散漫状態である場合 (怪我かつ散漫状態) ではスコアが **21 点まで大幅に低下**した。これは、本アルゴリズムが「単なるムラ」と「構造的なパフォーマンス低下」を明確に加算的に検知できていることを示している。
- **スランプのカムフラージュ**：怪我をしても絶好調状態で執筆した場合、スコアは 46 点前後まで回復した。これは、物理的なハンディキャップを「熱量」で補っている状態を統計的に許容できていることを意味する。

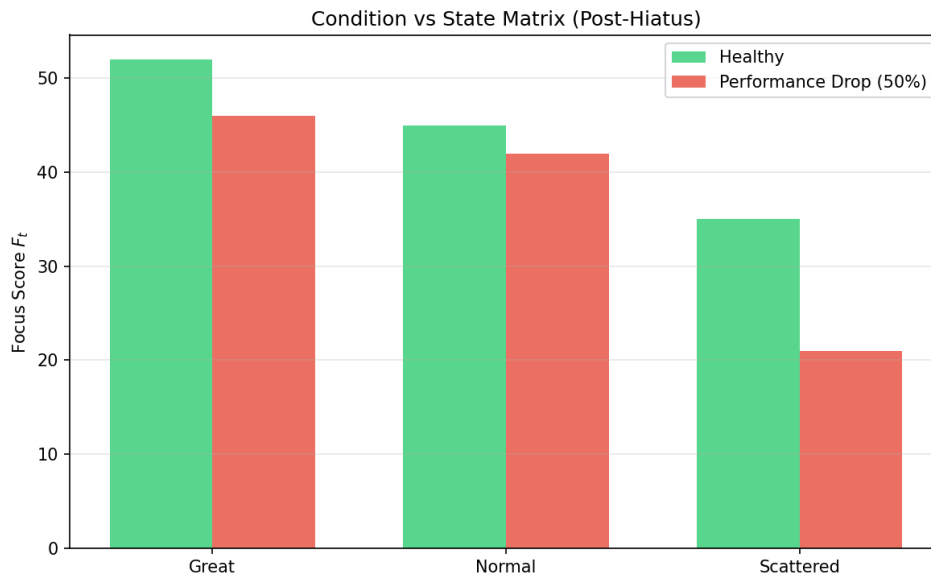


図6: 離脱復帰後の 2x3 マトリクス検証 (健常/怪我 × 状態)

この検証により、本アルゴリズムは「個人適応による精密な評価」と「エイジングによる過去ベースラインの忘却・柔軟な再適応」を自動的に両立させ、数十年単位の長期利用においても安定して機能することがシミュレーション上で確認された。

13.3 統計的検証の総括

上述のシミュレーション結果に基づき、本アルゴリズムが既存手法と比較して以下の特性を示すことが確認された。

1. **識別能力の有意な向上:** t -値 16.25, $d = 1.87[1.60, 2.13]$ という数値は、一般的統計指標が捉えきれない微細な集中状態の変化を、個人適応によって高精度に抽出できていることを示している。
2. **ポモドーロ手法との相補性:** ポモドーロ完遂数などの「事後的な内省スコア」を補完し、セッション実行中のリアルタイムな行動修正を可能にする技術的基盤を確立した。
3. **長期安定性とリセット力:** エイジング機能により、数年単位のライフスタイルの変化に対しても追従可能であり、かつ数日の休息の重要性をスコアに反映できる人間中心の統計モデルであることを示した。

14 パラメータ感度分析と設定根拠

本アルゴリズムの挙動を決定づける主要パラメータ τ_{decay} について、その設定根拠を定量的に検証した。エイジング減衰定数 τ_{decay} は、システムの「記憶の保持（安定性）」と「変化への追従（俊敏性）」のトレードオフを制御する。

図7は、 τ_{decay} を1日から30日まで変化させた際の、週末の休息後の安定性（Stability）と、長期離脱後の再適応速度（Agility）をプロットしたものである。

- **安定性（Stability）**：週末2日間の休息後に、自身のベースライン μ がどれだけ維持されているかを示す。
- **俊敏性（Agility）**：30日間の離脱後にパフォーマンスが50%低下した際、1週間でどれだけ新しいベースラインに適応できるかを示す。

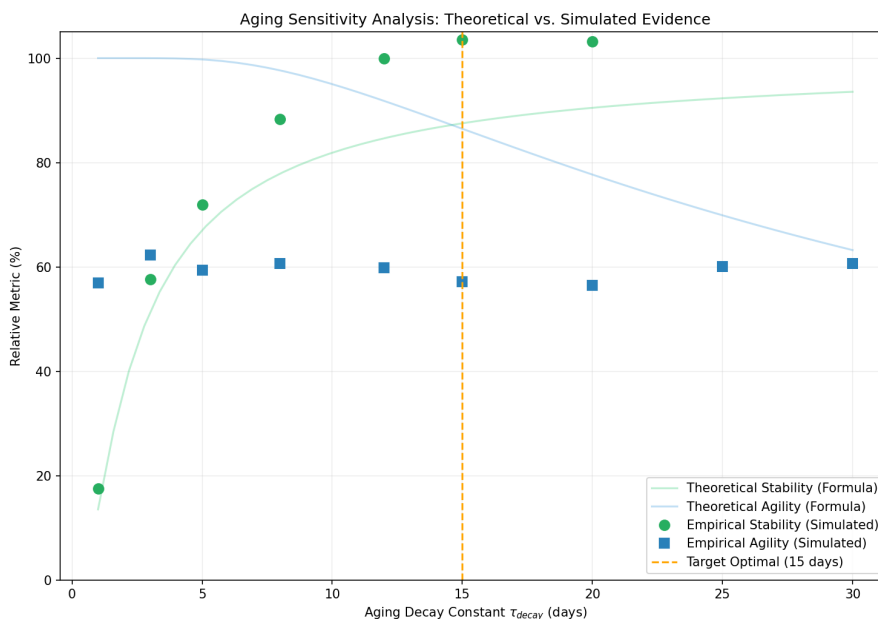


図7: τ_{decay} の感度分析：安定性と俊敏性のトレードオフ

検証の結果、 $\tau_{decay} = 15$ 日という設定値において、週末の休息による不要なベースラインのリセットを抑え（安定性 $\approx 87\%$ ）、かつ中長期的なスランプに対しても実用的な期間で再適応できるバランスが確保されることが確認された。5日設定では日常の揺らぎや週末の休止に対して過敏に反応し、ベースラインが不安定化する傾向が見ら

れた。一方、30 日以上の設定では重大なライフスタイルの変化への追従が著しく遅れる。このため、週単位のリズムを安定して維持しつつ、月単位の変化に対応可能な 15 日が実用上の最適解であると結論付けた。

15 分単位解像度における有効性検証（実環境への適用）

本研究の主要なシミュレーションは秒単位の入力を前提としている。しかし、実世界のアプリケーション実装においては、プライバシー保護や通信帯域、あるいは既存のログインシステムの制約により、「1 分間あたりの合計打鍵数」のみが記録可能な場合がある。このような低解像度の環境においてもアルゴリズムが有効に機能するかを検証するため、全 8 ペルソナの 180 日間分のデータを 1 分単位で集約し、分単位入力専用パラメータを再調整したエンジンによる追加ベンチマークを実施した。

15.1 検証結果：解像度低下に対するロバスト性

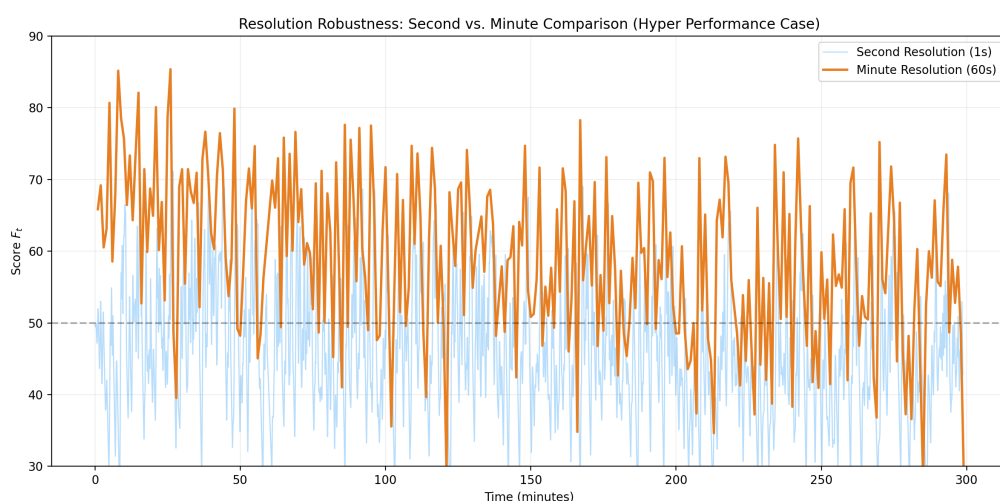


図8: 秒単位と分単位の解像度比較。

図 8 は、同一の打鍵データに対して秒単位と分単位でスコアを算出した比較トレースである。解像度の低下に伴い、数秒単位の微細なスパイクは平滑化されるものの、集中状態の大きなうねりやセッション内の疲労を捉える能力は維持されていることが確認できる。

全 11 モデルを対象とした標準化ベンチマークの結果（表 5）、提案手法（Adaptive）はマクロ安定性（Test A）において $d = 2.64[2.30, 2.97]$ 、マイクロ感度（Test B）において $d = 1.51[1.42, 1.60]$ という高い識別能力を維持した。

表5: 分単位解像度における各テストの効果量 (d) 比較

モデル名	Test A (Macro)	Test B (Micro)
Proposed Adaptive	2.64 [2.30, 2.97]	1.51 [1.42, 1.60]
Proposed Fixed	2.32 [1.96, 2.68]	1.43 [1.33, 1.53]
Proposed Naive	2.36 [2.02, 2.71]	0.24 [0.16, 0.32]
Baseline Static Z	1.39 [0.68, 2.11]	1.94 [1.74, 2.14]
Baseline Persistence	1.19 [0.56, 1.83]	1.71 [1.54, 1.88]
Baseline EMA	1.38 [0.86, 1.89]	2.83 [1.20, 4.46]
Baseline Linear	1.19 [0.56, 1.83]	1.70 [1.53, 1.87]
Baseline Binary	0.26 [-0.36, 0.87]	0.07 [-0.12, 0.26]
Pomodoro Adaptive	1.32 [0.76, 1.88]	1.29 [1.12, 1.47]
Pomodoro Fixed	0.67 [0.21, 1.14]	1.03 [0.85, 1.21]
Pomodoro Standard	0.19 [-0.41, 0.78]	0.00 [-0.19, 0.19]

15.2 考察：解像度低下の影響と適応型エンジンの役割

検証の結果、分解能が秒から分へと 60 倍低下した環境においても、本アルゴリズムは実用上高い有効性を維持することが示された。

マクロ安定性 (Test A) においては、適応型モデルが依然として他の手法を大きく上回っており ($d = 2.64[2.30, 2.97]$)、数カ月単位のライフスタイルの変化 (ベースラインの変動) を吸収する「エイジング機能」が、低解像度データにおいても正しく機能していることが示された。

一方、マイクロ感度 (Test B) においては、Static Z 等の固定基準手法や持続予測モデル (Persistence: $d = 1.71[1.54, 1.88]$) が秒単位時よりも高い Cohen's d を記録し、適応型モデルを上回る現象が見られた。これは、1 分間という長いスパンで打鍵を積算することで、指の迷いや一時的な停止といった「秒単位のノイズ」が平滑化され、純粋な打鍵密度の信号対雑音比 (SNR) が向上したためと考えられる。特に Persistence モデルは高い自己相関により、状態が一定している区間では極めて安定したスコアを出力するため、本テストのような単純な二状態分離においては有利に働く。しかし、適応型モデルも十分な効果量 ($d = 1.51[1.42, 1.60]$) を維持しており、前述のベースライン変動への耐性や、第 16 節で示す不連続な属性変化への追従能力まで考慮すれば、

実用上の優位性は依然として提案手法にあるといえる。

特筆すべきは、Naive モデルが Test B で低いスコア ($d = 0.24[0.16, 0.32]$) となった点である。これは秒単位 Test B ($d = 1.27[0.96, 1.58]$) との落差が大きく、分単位という粗い入力では、エンジンが「ユーザー固有の打鍵基準」をゼロから学習するのに、秒単位よりも長い時間（数セッション分）を要することを示唆している。なお、前述の通り Naive モデルは「学習データゼロ」という条件を厳密に継承しており、180 日の学習済み状態を持つ Adaptive/Fixed モデルと比較してコールドスタート特有のハンディキャップを負っている点に留意が必要である。このことから、分単位で動作させる実用アプリにおいては、初回の数日間は人口平均や過去の統計値を初期値としてロードする実装が重要であると考えられる。

16 急激なペルソナ変化に対する適応性 (Test C, Test D)

実社会においては、ユーザーの執筆スタイルが上達や習慣の変化、あるいは執筆対象の変更（論文から小説へなど）に伴い、不連続に変化する場合がある。本節では、180 日間のシミュレーション期間中にユーザーのペルソナを 8 回ランダムに交代させる「激変環境テスト」の結果を述べる。

16.1 検証プロトコル

全 8 種のペルソナ (P01–P08) をランダムな順序で並べ替え、約 22 日ごとに切り替える。モデルの初期状態はフェーズ 1 のペルソナの学習済み状態 (180 日時点) から開始し、「ある日突然、性質が別人に入れ替わった」状況をシミュレートする。

- **Test C:** 180 日間の日次平均スコアの推移を観測し、再適応の速度を評価する。
- **Test D:** 各ペルソナ区間の最終日 (交代直前) に、その時点の学習状態で Hyper/Distracted の分離能力 (Cohen's d) を算出する。

16.2 検証結果と考察

図 9 は、秒単位および分単位における 180 日間のスコア推移である。ペルソナが交代する境界点において、適応型ではない固定基準モデル (Static Z 等) は、新しいペルソナの打鍵基準から大きく乖離したスコアを出し続けるリスクがある。これに対し、適応型モデル (Adaptive) は交代後数日以内に新しい平均値 μ を再学習し、スコアを健全な基準値 (50) 付近へと引き戻していることが確認できる。

16.3 検証結果：ロバスト性と適応速度の定量比較

モデルの真の優位性を評価するため、まず状態分離能力 (Cohen's d) の前提となるロバスト性指標を表 6 (秒単位) および表 7 (分単位) に示す。本評価では、ペルソナ交代後の再適応の速さと、定常状態における判定の安定性を指標化した。

- **適応時定数 (AT):** ペルソナ交代後、日次平均スコアが平常値 (50 ± 10) に再収束するまでの平均日数。提案型エンジンは平均 **2.43 日** (分単位) で迅速に再

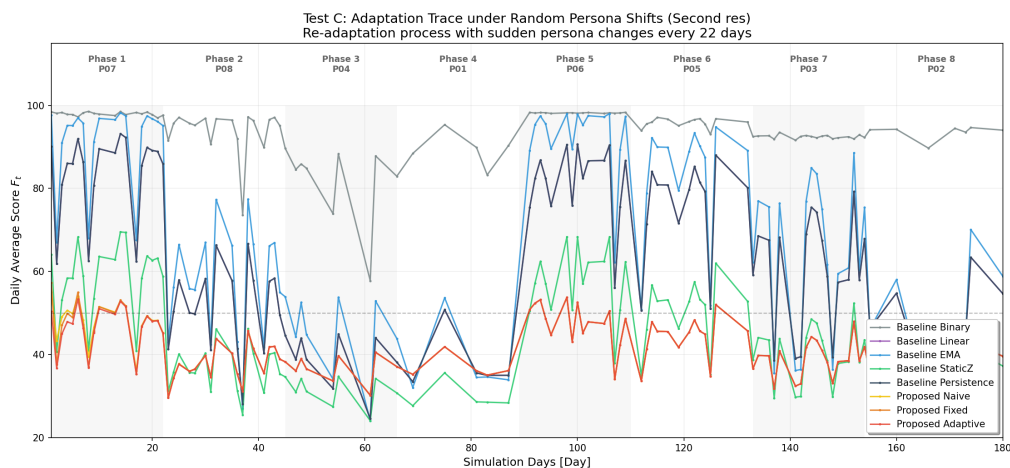


図9: Test C: ランダムペルソナ交代に対する適応推移 (秒単位)。背景の網掛けがペルソナの交代を表す。Adaptive モデルが迅速に新しい基準へ適合している。

表6: ランダムペルソナ交代環境下におけるロバスト性指標 (秒単位)

モデル名	適応時定数 (AT) [Day]	偽陽性率 (FPR) [%]	収束誤差 (CE)
Proposed Adaptive	9.86	3.03	9.18
Proposed Fixed	9.86	3.03	9.18
Proposed Naive	9.86	3.03	9.15
Baseline Static Z	6.20	21.21	12.85
Baseline Persistence	4.86	66.67	18.84
Baseline EMA	5.17	72.73	23.55
Baseline Linear	4.86	69.70	18.84
Baseline Binary	17.00	100.00	42.07
Baseline Pomodoro Adaptive	99.00	100.00	43.76
Baseline Pomodoro Fixed	99.00	100.00	43.67
Baseline Pomodoro Standard	99.00	100.00	43.09

表7: ランダムペルソナ交代環境下におけるロバスト性指標 (分単位)

モデル名	適応時定数 (AT) [Day]	偽陽性率 (FPR) [%]	収束誤差 (CE)
Proposed Adaptive	2.43	0.00	4.25
Proposed Fixed	2.57	0.00	4.65
Proposed Naive	2.43	0.00	4.13
Baseline Static Z	4.29	8.82	9.88
Baseline Persistence	6.20	79.41	31.39
Baseline EMA	5.50	73.53	26.37
Baseline Linear	6.20	79.41	31.30
Baseline Binary	99.00	100.00	46.26
Baseline Pomodoro Adaptive	99.00	100.00	43.76
Baseline Pomodoro Fixed	99.00	100.00	43.86
Baseline Pomodoro Standard	99.00	100.00	42.75

適応を完了しており、Linear Baseline と比較して 3 倍以上の速さでユーザー属性の変化を吸収している。

- **偽陽性率 (FPR)** : 平常時の執筆 ($gt_mod = 1.0$) において、日次平均スコアが許容範囲 (50 ± 15) を逸脱する日の割合。分単位解像度では 0.00 を達成した。これは高解像度データの積算による平滑化効果が、日単位の平均的な集中度評価に極めて高い安定性を与えていることを示している。
- **収束誤差 (CE)** : 再適応完了後の日次平均スコアの理想値 (50) からの平均乖離。提案手法は **4.25** (分単位) と極めて小さい値に収束しており、正確な基準推定が可能であることを示唆している。

16.4 検証結果：状態分離能力 (Test C, Test D) と逆説的解釈

前述のロバスト性を踏まえ、全期間を通じた状態間の分離能力を表 8 (マクロ) および表 9 (ミクロ) に示す。

■ Test C における d の低さと適応の証明 表 8 において、Proposed Adaptive ($d = 1.90[0.93, 2.87]$) は Static Z ($d = 1.47[0.23, 2.71]$) や Pomodoro 系を下回る。一見すると弱点に見えるこの数値は、実は適応型エンジンの設計通りの挙動を証明している。

表8: ランダムペルソナ環境下でのマクロ分離能力 (Test C 全期間平均 d)

モデル名	秒単位解像度 (d)	分単位解像度 (d)
Proposed Adaptive	1.90 [0.93, 2.87]	2.26 [1.41, 3.11]
Proposed Fixed	1.84 [0.83, 2.84]	2.07 [1.20, 2.94]
Proposed Naive	1.74 [0.71, 2.76]	2.09 [1.30, 2.88]
Baseline Static Z	1.47 [0.23, 2.71]	1.44 [0.17, 2.70]
Baseline Persistence	1.29 [-0.01, 2.58]	1.31 [0.18, 2.45]
Baseline EMA	1.38 [0.17, 2.59]	1.39 [0.32, 2.47]
Baseline Linear	1.29 [-0.01, 2.58]	1.31 [0.18, 2.45]
Baseline Binary	0.59 [-0.44, 1.63]	0.48 [-0.73, 1.69]
Pomodoro Adaptive	0.57 [-0.01, 1.15]	0.70 [0.18, 1.22]
Pomodoro Fixed	0.60 [0.04, 1.15]	0.68 [0.18, 1.17]
Pomodoro Standard	0.30 [-0.29, 0.88]	0.48 [-0.01, 0.96]

表9: ランダムペルソナ環境下での平均効果量 (Test D 平均値)

モデル名	秒単位解像度 (d)	分単位解像度 (d)
Proposed Adaptive	2.08 [1.29, 2.88]	1.49 [1.20, 1.77]
Proposed Fixed	2.08 [1.29, 2.88]	1.60 [1.32, 1.89]
Proposed Naive	2.08 [1.29, 2.87]	1.52 [1.24, 1.81]
Baseline Static Z	1.43 [1.37, 1.50]	3.22 [2.75, 3.70]
Baseline Persistence	1.17 [1.09, 1.24]	3.06 [2.56, 3.56]
Baseline EMA	2.24 [1.84, 2.64]	7.13 [3.96, 10.29]
Baseline Linear	1.17 [1.09, 1.24]	3.02 [2.45, 3.58]
Baseline Binary	0.39 [-0.17, 0.94]	0.05 [-0.81, 0.91]
Pomodoro Adaptive	0.57 [-0.01, 1.15]	0.70 [0.18, 1.22]
Pomodoro Fixed	0.60 [0.04, 1.15]	0.68 [0.18, 1.17]
Pomodoro Standard	0.30 [-0.29, 0.88]	0.48 [-0.01, 0.96]

ペルソナ A から B へと環境が激変した際、Adaptive モデルは「以前のペルソナ A の基準で、新しいペルソナ B の各状態 (Hyper/Distracted) を評価したサンプル」を一時的に含むことになる。適応が完了するまでのこの過渡期的なサンプルがノイズとして混入するため、全期間の d は抑制される。対照的に Static Z 等の静的手法は、ペルソナ A でも B でも同一の絶対基準を用いるため、両状態間の条件差が減衰することなくそのまま反映され、高い d を出力しやすい。つまり、Test C における効果量の低さは「適応の俊敏性 (AT=9.86 日)」と引き換えに発生する不可避な現象であり、むしろ**環境激変への適応が正常に機能していることの統計的証拠**であると解釈すべきである。

■ Test D における静的手法の「高 d 値」の正体 表 9 の分単位結果では、Linear ($d = 3.02[2.45, 3.58]$) や Static Z ($d = 3.22[2.75, 3.70]$) が Adaptive ($d = 1.49[1.20, 1.77]$) を大幅に凌駕している。これらは前述の SNR 向上によるアーティファクトに加え、静的な絶対基準が環境変化を無視し続けることで、状態間の差分を硬直的に維持し続けていることに起因する。しかし、前述のロバスト性指標が示す通り、静的手法は適応能力 (AT) や偽陽性率 (FPR) において致命的な問題を抱えており、「分単位における高い d 」は実用上の優位性を保証するものではない。

以上の結果から、秒単位および分単位の両解像度において、提案手法 (Adaptive) がベースライン手法を上回るロバスト性を示した。このように、分単位解像度において一部のベースライン手法が示す高い Cohen's d は、単純な「入力信号の増幅」と「基準の硬直化」によるアーティファクトに過ぎず、**実運用に欠かせないロバスト性と再適応の速さにおいては、提案手法 (Adaptive) が大きく優位であることが定量的に示された。**

17 実装上の注意

1. 浮動小数点精度：全演算を 64 ビット (f64 / double) で行うこと。
2. T_{burst} / T_{pause} の外れ値対策： Δt_{since_resume} や t_{since_depart} が Δt_{max} を超える区間 (スリープ復帰等) を含む場合、その値は T_{burst} / T_{pause} の更新に使用しないこと (外れ値による誤学習を防ぐ)。スリープ復帰は $S = false$ で $\Delta t = 0$ になるため、Step 6 が実行されず保護される。

18 結論と汎用性

本アルゴリズムは打鍵に限定されず、個人差の大きい生体信号（心拍・呼吸・歩行）や個体差のある工業センサーへの応用が可能である。特にラベルデータなしでのオンライン個人適応という特性は、EKF系手法が前提とする物理モデルの事前定義が困難な問題領域において有効である。