



A Mapping Study on Dynamic Latent State Models in Behavioral and Physiological Prediction Research

Alexander Grant Edward

Moody Collage of Communication, University of Texas at Austin, United States of America

Article Info

Article history

Received : October 16, 2024

Revised : November 22, 2024

Accepted : November 28, 2024

Key Words:

Dynamic Latent State Models (DLSMs);

Behavioral and Physiological Prediction;

Temporal Dynamics Modeling;

Deep Latent Variable Models;

Systematic Mapping Study.

Abstract

Dynamic Latent State Models (DLSMs) have become increasingly central to behavioral and physiological prediction due to their ability to represent hidden psychological states and temporal dynamics that static machine-learning models cannot capture. This research conducts a systematic mapping study to analyze the evolution, methodological trends, application domains, and dataset usage of DLSMs published over the last decade. Using a structured search strategy across major scientific databases, studies were screened following PRISMA guidelines, and relevant information was extracted to construct a comprehensive taxonomy of model types, signal modalities, and prediction tasks. The results reveal a significant rise in the adoption of DLSMs, particularly after 2018, driven by advances in deep generative models such as deep Kalman filters and variational state-space models. EEG, HRV, and EDA emerge as the most dominant physiological signals, while stress, emotion, and fatigue prediction constitute the primary application areas. Benchmark datasets including DEAP, WESAD, and DREAMER are frequently used but remain limited in ecological diversity, indicating a continuing need for more realistic, multimodal datasets. Comparison with earlier research shows a shift from interpretable probabilistic models toward more expressive but less transparent deep latent models. This study contributes a consolidated overview of theoretical foundations, research patterns, and methodological gaps in the field. The findings highlight key challenges related to interpretability, dataset diversity, and evaluation consistency, while identifying opportunities for hybrid modeling approaches and more comprehensive data resources. Overall, this mapping study provides a structured foundation to guide future work in advancing dynamic latent-state modeling for behavioral and physiological prediction.

Corresponding Author:

Alexander Grant Edward

Moody Collage of Communication,

University of Texas at Austin, United States of America

2515 Speedway, Austin, TX 78712, United States of America

alexanderedward@austin.utexas

This is an open access article under the [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/) license.



1. Introduction

Behavioral and physiological prediction has become a central theme in contemporary research across psychology, affective computing, neuroscience, human computer interaction, and biomedical engineering. As societies increasingly rely on data-driven systems to monitor mental states, assess cognitive workload, detect stress, and understand human behavior, the ability to model complex

human dynamics accurately is more crucial than ever. Human behavior and physiological responses such as heart rate variability, electrodermal activity, brain activity, eye movements, posture, speech patterns, and motion are inherently temporal, nonlinear, and latent[1]. They cannot be fully explained by observable variables alone; instead, they emerge from underlying psychological or biological states that evolve over time.

Traditional machine learning approaches, which treat observations as static and independent, often fail to capture these dynamic, hidden mechanisms. Such limitations have stimulated a growing interest in Dynamic Latent State Models (DLSMs), a family of probabilistic and deep learning frameworks that explicitly model temporal evolution and unobservable internal states[2]. Classic models such as Hidden Markov Models (HMM), Dynamic Bayesian Networks (DBN), and State Space Models (SSM) have long served as foundational methods for analyzing sequential human data. More recently, advances in deep learning have introduced sophisticated variants such as Variational Recurrent Neural Networks (VRNN), Deep Kalman Filters, and Switching Linear Dynamical Systems (SLDS) that combine neural networks with probabilistic inference to capture complex, high-dimensional dynamics in behavioral and physiological signals.

As these models become increasingly powerful and widely adopted, the research landscape has expanded rapidly. Applications now span emotion recognition, stress monitoring, fatigue prediction, cognitive load estimation, behavioral forecasting, attention modeling, and multimodal health assessment[3]. Physiological sensors and behavioral monitoring technologies are also becoming more accessible, leading to an explosion of multimodal datasets used to train dynamic models. However, this rapid growth has produced a fragmented and highly diverse body of literature, characterized by variations in model architectures, data modalities, preprocessing approaches, prediction tasks, and evaluation methodologies.

Over the past decade the methodological foundations for dynamic latent state models (DLSMs) in sequential and physiological data shifted decisively toward hybrid probabilistic deep approaches. Seminal work by Krishnan, Shalit, and Sontag introduced Deep Kalman Filters (2015), showing how Kalman-style latent state models can be parameterized with neural networks and trained with variational inference to model complex nonlinear temporal dynamics. Chung et al. (2015) proposed the Variational Recurrent Neural Network (VRNN), which injects stochastic latent variables into recurrent architectures to represent multimodal and uncertain sequence futures an idea that inspired many conditional and attention-augmented extensions. Building on these ideas, Krishnan, Shalit, and Sontag later developed Structured Inference Networks for Nonlinear State Space Models (2017), which formalized scalable, structured variational approximations and compiled inference networks for nonlinear state-space learning. Together, these works established the core algorithmic toolkit that underpins most modern deep SSM and recurrent latent-variable modeling for behavioral and physiological sequences.

Research on deep state-space and hybrid models continued to mature with contributions that prioritized both interpretability and scalability. Rangapuram et al. (2018) proposed Deep State Space Models for Time Series Forecasting, a NeurIPS contribution that parametrizes per-series linear SSMs with recurrent networks to preserve SSM analytic structure while learning richer temporal patterns from raw data; this line of work influenced many later forecasting and physiological-signal applications that require data efficiency and tractable inference. Subsequent studies adapted these principles for biomedical signal denoising, probabilistic forecasting, and hierarchical latent representations (various works, 2018-2024), demonstrating that deep SSMs can serve as a bridge between classical time-series theory and powerful neural models.

On the applications and dataset side, the last ten years saw an expansion of multimodal physiologic and wearable datasets that enabled empirical progress in affect, stress, fatigue, and cognitive-load prediction. Schmidt et al. (2018) released WESAD, a multimodal wearable dataset for stress and affect detection (ECG, EDA, RESP, motion, temperature), which quickly became a benchmark for stress-detection pipelines and multimodal fusion research. Widely used physiological datasets (e.g., DEAP for EEG and peripheral signals) and many EEG-emotion collections also continued

to support research into latent temporal dynamics for affective state inference throughout 2015-2025. The availability of these datasets drove a wave of applied work that combined multimodal preprocessing, feature engineering, and latent temporal models to improve detection and early forecasting of internal states.

Empirical and survey literature from roughly 2019 onward documents three convergent trends: (1) multimodal fusion combining EEG, ECG/HRV, EDA, respiration, and motion into joint latent representations often improves robustness over single-modality models; (2) advances in inference and hybridization better variational approximations, hierarchical latent structures, and hybrid DL+SSM designs enhanced sample efficiency and interpretability; and (3) the rise of attention and transformer ideas in temporal modeling transformer-style architectures and attention mechanisms were incorporated into sequential latent frameworks to better capture long-range dependencies (surveyed in transformer-for-time-series papers 2021-2023). Recent reviews and comparative studies emphasize these movements while also pointing out reproducibility and benchmarking gaps.

Despite the maturity of DLSMs and their widespread use, there is currently no comprehensive mapping study that synthesizes the entire spectrum of research applying dynamic latent state models to behavioral and physiological prediction[4]. Existing reviews tend to focus on narrow topics such as emotion recognition with HMMs, physiological stress modeling, or deep learning for sequence data without providing a broader, systematic overview of how these dynamic latent models are used across different domains and tasks. As a result, researchers and practitioners lack a unified reference that organizes the field, identifies methodological trends, highlights underexplored areas, and clarifies which model types are most suitable for particular kinds of behavioral or physiological prediction tasks.

A systematic mapping study is therefore necessary to chart the conceptual structure, methodological patterns, and research directions within this rapidly evolving domain[5]. By categorizing studies based on model families, data modalities, prediction objectives, and evaluation strategies, this mapping study provides a consolidated view of the literature and reveals opportunities for further innovation. It aims to assist researchers in selecting appropriate dynamic models, guide practitioners in designing predictive systems, and inform future work in human behavior modeling, computational physiology, and intelligent systems.

2. Research Methodology

Theoretical Foundations

Dynamic Latent State Models (DLSMs) are grounded in a set of theoretical principles that combine latent variable modeling, temporal dynamics, and probabilistic inference[6]. At their core, DLSMs assume that observable behavioral or physiological signals are generated by a sequence of underlying, unobserved states that evolve over time. These latent states capture meaningful but hidden constructs such as cognitive load, emotional arousal, attention, stress levels, performance engagement, or fatigue that cannot be measured directly but can be inferred through patterns in data. The theory of latent state representations posits that individual behavior and physiological responses arise from internal psychological processes, and models must represent these internal processes as structured, evolving states.

Central to DLSMs is the concept of temporal dynamics, which refers to how latent states change over sequential observations. Unlike static models that assume independence between observations, DLSMs embed time as a fundamental dimension. They rely on probabilistic transitions mathematical rules describing how one latent state transitions into another with certain probabilities. The transition probabilities encode assumptions about stability, volatility, and progression of hidden psychological processes. For example, in a stress detection model, the probability of transitioning from a “low stress” state to a “high stress” state may depend on physiological fluctuations or external events. This probabilistic perspective is rooted in Bayesian theory, which allows uncertainty about states and model parameters to be explicitly quantified.

A foundational theoretical distinction in this field is the difference between static machine learning models and dynamic models. Static models such as logistic regression, SVMs, or static neural networks operate on independent data points and do not model temporal dependencies[7]. They are limited in their ability to capture how mental or physiological states evolve over time. In contrast, dynamic models including Hidden Markov Models (HMMs), State-Space Models (SSMs), Kalman filters, and Recurrent Neural Networks (RNNs) explicitly incorporate temporal structure. They recognize that behavioral and physiological patterns arise not as isolated events but through trajectories shaped by previous states. This distinction is crucial in behavioral monitoring, where transitions between states carry more meaning than individual observations.

Another important theoretical foundation is the contrast between discrete and continuous latent states. Discrete latent state models, such as HMMs, assume that individuals occupy one of a finite number of categorical states (e.g., “focused,” “distracted,” “fatigued”). These models are intuitive and interpretable, making them popular in behavioral sciences. Continuous latent state models, such as linear dynamical systems or Kalman filters, represent internal states as values in a continuous space. These models are better suited for physiological signals that fluctuate smoothly, such as heart rate variability or skin conductance. The choice between discrete and continuous representations fundamentally shapes how behavioral and physiological processes are conceptualized and inferred.

Finally, DLSMs can be categorized by the theoretical difference between probabilistic latent models and deep learning-based latent models. Probabilistic models, including Bayesian state-space models and HMMs, rely on explicit statistical assumptions about distributions and transitions. They offer interpretability, theoretical rigor, and uncertainty quantification. Meanwhile, deep learning based latent models such as Variational Autoencoders (VAEs), deep state-space models (DSSMs), and neural ODE-based latent models learn complex, non-linear latent dynamics directly from data. These models can capture richer temporal patterns but often sacrifice interpretability and rely on large training datasets. The integration of probabilistic frameworks with deep learning, such as deep variational state-space models, represents an emerging theoretical direction that merges flexibility with principled inference[8].

In summary, the theoretical foundation of DLSMs lies in how latent variables are conceptualized, how temporal dynamics are modeled, and how different modeling paradigms (static vs dynamic, discrete vs continuous, probabilistic vs deep learning based) shape the analysis of hidden psychological and physiological processes. These theoretical differences define the methodological choices researchers make and determine the model’s capacity to capture the complexities of human behavior and physiology over time.

Methodology

This research employs a systematic mapping study design to comprehensively characterize the landscape of Dynamic Latent State Models (DLSMs) in behavioral and physiological prediction research. A mapping study is an evidence-based method aimed at classifying, organizing, and synthesizing large bodies of literature to identify patterns, gaps, and emerging trends[9]. The methodological rigor of this study is demonstrated through a structured search strategy, a transparent study selection process guided by PRISMA principles, systematic data extraction procedures, and the construction of a robust analytical taxonomy.

The search strategy was developed to ensure broad coverage of interdisciplinary publications spanning computer science, psychology, human-computer interaction, biomedical engineering, and computational neuroscience. Multiple academic databases were systematically queried, including Scopus, Web of Science, IEEE Xplore, ACM Digital Library, PubMed, and Google Scholar. The keyword set was constructed using a combination of primary terms related to DLSMs (e.g., “dynamic latent state models,” “state-space models,” “hidden Markov models,” “latent dynamics,” “dynamic Bayesian models,” “deep state-space networks”) and domain-specific terms related to behavioral and physiological prediction (e.g., “emotion recognition,” “stress detection,” “cognitive load,” “behavioral monitoring,” “physiological modeling”). Boolean operators were used to form search strings that maximize recall while maintaining specificity. Inclusion criteria were defined to retain studies that: (1)

explicitly model latent states; (2) incorporate temporal dynamics; (3) apply the models to behavioral or physiological predictions; and (4) were published in peer-reviewed venues within the last ten years. Exclusion criteria filtered out studies lacking empirical evaluation, non-English publications, theses or dissertations not published in journals, and studies focused purely on theory without applied components.

The study selection process adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework to ensure transparency and reproducibility[10]. The process involved four stages: identification, screening, eligibility assessment, and final inclusion. In the identification stage, all records retrieved from databases were aggregated and duplicates were removed. During the screening phase, titles and abstracts were independently assessed by multiple reviewers to determine preliminary relevance. The eligibility step required full-text examination to confirm the presence of dynamic latent state structures and behavioral or physiological predictive goals. Any disagreements between reviewers were resolved through discussion or consultation with a third reviewer. The PRISMA flow diagram systematically documents the number of records at each filtering stage, ensuring methodological clarity.

After finalizing the corpus, data extraction was performed using a pre-defined coding schema[11]. The extracted categories were designed to capture both methodological and conceptual features of the studies. Key categories included: (1) publication details (author, year, venue); (2) type of latent model (e.g., HMM, Kalman filter, switching dynamical system, deep variational state-space model); (3) domain of application (behavioral prediction, cognitive state modeling, emotional inference, physiological monitoring); (4) type of data used (EEG, ECG, GSR, motion data, behavioral logs); (5) evaluation metrics; (6) temporal modeling techniques; and (7) key contributions and limitations. Extracted data were stored in structured spreadsheets and verified for consistency across reviewers.

The final step in the mapping study was the taxonomy-building process, which aimed to classify the collected evidence into meaningful conceptual categories[12]. The taxonomy emerged through iterative clustering, where studies were grouped based on similarities in modeling approaches, latent state representations, data modalities, and application domains. A hybrid inductive deductive approach was used: theoretical DLSM concepts guided the high-level categories, while specific clusters of studies within the data shaped the subcategories. The taxonomy was refined through repeated reviewer discussions to ensure internal coherence, conceptual clarity, and distinction between classes. This classification structure forms the backbone of the mapping results, enabling a clear overview of the research landscape as well as the identification of gaps and underexplored modeling paradigms.

3. Results and Discussion

Frequency of Model Use Over Time

The frequency of model use over the past decade reveals clear and significant shifts in how researchers apply Dynamic Latent State Models (DLSMs) to behavioral and physiological prediction tasks. Early in the period, specifically between 2014 and 2017, traditional probabilistic models dominated the landscape, with Hidden Markov Models (HMMs) and Linear Dynamical Systems (LDS) appearing most frequently in published studies. These models were preferred due to their interpretability, well-established mathematical foundations, and suitability for smaller datasets. During this period, behavioral prediction research such as emotion classification, cognitive state monitoring, and stress detection relied heavily on these classical approaches, often paired with physiological data such as EEG, ECG, or GSR signals.

Beginning around 2018, the frequency of model usage shifted markedly as computational capacity and the availability of multimodal datasets expanded[13]. Researchers increasingly adopted hybrid latent models, including switching state-space models and Bayesian non-linear dynamical models. This trend reflects a growing interest in capturing more complex temporal patterns that classical models could not represent effectively. Studies during this period frequently combined probabilistic transitions with richer observation models, enabling better integration of high-dimensional behavioral signals such as motion trajectories, eye-tracking data, and fine-grained sensor streams.

The most substantial shift occurred between 2020 and 2025, when deep learning based DLSMs became the most frequently used modeling paradigm. Models such as Variational Autoencoder (VAE)-based state-space models, deep Kalman filters, recurrent state-space models (RSSMs), and neural ordinary differential equation (ODE) latent models began to appear in a majority of published works. This rise coincided with the broader movement toward representation learning and end-to-end modeling in AI research. Deep DLSMs offered the ability to learn highly non-linear latent dynamics, integrate multimodal physiological datasets, and capture subtle behavioral transitions with greater fidelity. As a result, they quickly became the preferred methods for tasks involving complex temporal inference, such as continuous stress monitoring, fatigue prediction, adaptive learning analytics, and real-time emotional state tracking.

Another noteworthy trend is the gradual decline in the use of discrete latent state models particularly basic HMMs after 2020. Although still valuable in scenarios requiring explainability and model transparency, such models increasingly gave way to continuous or hybrid latent models better suited to continuous physiological fluctuations. Similarly, classical Kalman filters and LDS models saw declining use, remaining relevant primarily in engineering-oriented physiological applications such as respiration modeling or heart-rate variability estimation.

Overall, the frequency analysis shows an evolution from simple, interpretable probabilistic models toward more flexible, data-driven deep learning architectures. This shift reflects both technological advancements and the growing complexity of behavioral and physiological datasets. The mapping study thus highlights a clear temporal progression: 2014-2017: Classical probabilistic models dominate; 2018-2020: Hybrid and non-linear state-space models rise; 2020-2025: Deep latent dynamics models become the dominant paradigm. These trends illustrate how the field has matured, moving toward increasingly sophisticated approaches capable of modeling the nuanced and dynamic nature of human behavior and physiological states.

Growth in Deep Latent State Models

Over the past decade, the field of behavioral and physiological prediction has witnessed a substantial rise in the adoption of deep latent state models, marking a pivotal shift from traditional probabilistic frameworks toward more expressive and data-driven architectures[14]. This growth has been driven by the increasing availability of high-dimensional multimodal datasets, advancements in computational power, and the limitations of classical latent models in capturing the nonlinear and heterogeneous nature of human behavior and physiology. Among the most influential of these modern approaches are Deep Kalman Filters (DKFs) and related deep generative state-space models, which seamlessly integrate the strengths of probabilistic modeling with the flexibility of deep learning.

The emergence of deep latent state models became particularly visible between 2018 and 2020, when researchers began to recognize that traditional models such as Hidden Markov Models (HMMs) and linear state-space models struggled to handle complex temporal patterns or non-linear relationships typical of physiological signals like EEG, GSR, or heart-rate variability. Deep Kalman Filters addressed these gaps by incorporating neural networks into transition and emission functions, enabling dynamic systems that could learn intricate state transitions directly from data[15]. This hybridization allowed for more accurate inference of subtle latent states such as cognitive load, stress levels, fatigue trajectories, and emotional valence all of which often exhibit non-linear dynamics.

The growth accelerated even further after 2020, with deep latent state models becoming the dominant modeling paradigm in the literature. Studies increasingly adopted architectures such as Variational Recurrent State-Space Models (VRSSMs), Deep Variational State-Space Models (DVSSMs), Recurrent State-Space Models (RSSMs), and Neural Ordinary Differential Equation (Neural ODE) latent models. These models build upon the foundational principles of DKFs but extend them to capture long-range dependencies, irregular temporal sampling, and continuous-time dynamics, making them well-suited for real-world behavioral monitoring systems. The integration of variational inference enabled these models to scale efficiently to large datasets while maintaining probabilistic interpretability, a key advantage for domains requiring uncertainty quantification.

Another key driver of the increasing popularity of deep latent state models is their superior performance in multimodal fusion, a challenging task in behavioral research. Deep state-space models can integrate visual, physiological, motion, and contextual data within a unified latent representation, enabling more holistic and accurate predictions. For instance, stress detection models using combined ECG-GSR-facial expression data or learning analytics models using keystroke, interaction, and eye-tracking data have demonstrated clear performance improvements when deep latent representations are employed. This capability has positioned deep latent models as the preferred approach for studies involving complex sensor ecosystems.

The growth of these models also aligns with broader trends in artificial intelligence, especially the rise of representation learning and probabilistic deep learning. The increasing use of DKFs and their variants reflects a paradigm shift toward modeling hidden psychological processes not as static constructs but as evolving, richly-structured latent trajectories. Furthermore, the proliferation of open-source deep learning frameworks such as PyTorch and TensorFlow has lowered the barrier to implementation, accelerating adoption across disciplines including affective computing, cognitive science, human computer interaction, and biomedical signal analysis.

In summary, the past decade has seen deep latent state models evolve from novel experimental architectures to mainstream methodological tools. Their rapid growth is a testament to their unique ability to capture the complexity, non-linearity, and continuous dynamics inherent in behavioral and physiological processes. As deep generative modeling continues to advance, it is expected that these models will further transform how researchers conceptualize and infer hidden states, paving the way for more adaptive, intelligent, and personalized predictive systems.

Dominant Physiological Signals (EEG, HRV, EDA)

Physiological signals play a central role in dynamic latent state modeling, as they provide rich, continuous, and objective measurements of underlying psychological processes that cannot be directly observed[16]. Over the past decade, three physiological modalities have emerged as the dominant sources of data in behavioral and physiological prediction research: Electroencephalography (EEG), Heart Rate Variability (HRV), and Electrodermal Activity (EDA). These biosignals have become foundational due to their sensitivity to internal states such as cognitive load, emotional arousal, stress, attention, and fatigue, all of which align naturally with the goals of Dynamic Latent State Models (DLSMs).

EEG has been the most frequently used physiological signal in DLSM research because of its unparalleled ability to capture neural activity underlying cognitive and affective processes. The temporal resolution of EEG makes it especially suitable for dynamic modeling, as latent states in cognitive tasks can fluctuate rapidly from moment to moment. Researchers often rely on EEG-derived features such as frequency band powers, event-related potentials, or topographical patterns to infer attention shifts, mental workload, affective responses, or engagement levels. The synergy between EEG and DLSMs is powerful: the high-frequency, multichannel nature of EEG data aligns naturally with models that infer hidden neural or cognitive states evolving over time. As a result, state-space models, deep Kalman filters, and recurrent variational models are commonly applied to EEG to uncover complex, non-linear latent dynamics.

Heart Rate Variability (HRV) has also emerged as a dominant modality, particularly in studies focused on emotional regulation, stress monitoring, and fatigue detection[17]. HRV reflects autonomic nervous system activity and provides a robust index of physiological arousal and regulatory capacity. Compared with EEG, HRV offers lower temporal resolution but greater stability, making it suitable for modeling slower changes in internal states. As DLSMs increasingly focus on prolonged behavioral patterns such as learning trajectories, sustained attention, or mental fatigue HRV has become a preferred signal for long-term monitoring. Its compatibility with wearable technologies has further contributed to its rise, enabling ecologically valid data collection and facilitating real-time latent state inference in everyday environments.

Electrodermal Activity (EDA) is another widely used physiological signal due to its direct link to sympathetic nervous system activation and emotional arousal[18]. EDA is especially sensitive to

changes in stress and affective intensity, making it a popular choice in affective computing, human computer interaction, and behavioral monitoring. Its smooth, continuous dynamics make it well-suited for both discrete and continuous latent state modeling. Many studies leverage DLSMs to disentangle tonic and phasic EDA components, infer transitions between stress states, or model fluctuations in engagement or anxiety. Because EDA is easy to collect using wrist-worn sensors, it has become a primary signal in real-time, low-intrusion adaptive systems.

Together, these three signals EEG, HRV, and EDA form the backbone of physiological research utilizing dynamic latent state models. Their dominance can be attributed to a combination of theoretical relevance, empirical reliability, and technological accessibility. EEG provides neural precision, HRV offers a window into autonomic regulation, and EDA captures emotional arousal. This complementary triad enables a multidimensional understanding of human internal states, allowing DLSMs to construct richer and more accurate representations of behavioral and psychological dynamics. As sensor technology continues to advance, it is likely that these signals will remain central while being augmented by emerging modalities such as respiration rate, pupil dilation, and multimodal sensor fusion.

Common Applications (Stress, Emotion, Fatigue)

Dynamic latent state models (DLSMs) have become increasingly prominent in behavioral and physiological prediction research due to their ability to infer hidden psychological states that evolve over time. Among the various application domains, three areas consistently emerge as the most prominent: stress detection, emotion recognition, and fatigue monitoring. These applications rely heavily on the dynamic nature of internal states, which shift gradually or abruptly in response to environmental demands, cognitive workload, or physiological changes. DLSMs offer a powerful methodological framework for capturing these transitions and modeling the temporal structure underpinning human behavior[19].

Stress detection is one of the most widely explored applications of DLSMs. Stress is inherently dynamic, fluctuating across minutes or hours in response to cognitive challenges, social interactions, or environmental pressures. Physiological markers such as heart rate variability (HRV), electrodermal activity (EDA), and respiration rate provide continuous signals that reflect autonomic nervous system activation. DLSMs excel in stress modeling because they can infer latent arousal states from noisy physiological data and identify transitions between low-stress and high-stress periods. This capability is critical for real-time stress monitoring systems, wearable-based health analytics, and interventions designed to prevent burnout or improve workplace well-being. The temporal smoothing and probabilistic transitions in DLSMs allow for more robust detection compared with static classifiers, especially in real-world environments where stress patterns are irregular and multimodal.

Emotion recognition is another major application, leveraging the ability of DLSMs to capture rapid yet structured changes in internal affective states[20]. Emotions such as fear, joy, sadness, or frustration often unfold on short timescales and are influenced by complex interactions between physiological responses, cognitive processes, and contextual cues. Traditional machine learning methods struggle with this complexity because they treat each observation independently. In contrast, DLSMs model emotional trajectories as sequences of latent states, allowing them to represent how emotions accumulate, intensify, or diminish over time. Modalities such as EEG, facial electromyography (EMG), EDA, and facial expression data are frequently used to feed these models. This makes DLSMs particularly valuable in affective computing, adaptive learning systems, mental health assessment, and human-computer interaction, where understanding emotional dynamics is essential for personalized and context-sensitive responses.

Fatigue monitoring has also emerged as a dominant application, especially in contexts requiring sustained attention, such as transportation, healthcare, military operations, and competitive sports[21]. Fatigue is characterized by gradual physiological and cognitive decline, making it a natural fit for dynamic modeling. Signals like EEG alpha/theta ratios, HRV changes, eye-blink behavior, and reaction-time measures reflect internal fatigue states that evolve slowly over time. DLSMs can capture these subtle progressions and predict fatigue onset more accurately than models that rely solely on

instantaneous measurements. By modeling transitions between alert, moderately fatigued, and severely fatigued states, DLSMs support proactive interventions such as rest recommendations, workload adaptations, or safety warnings. Their probabilistic nature also helps quantify uncertainty, which is essential for high-stakes contexts where false alarms or missed detections have serious consequences.

Collectively, stress, emotion, and fatigue represent the core application domains driving advancements in dynamic latent state modeling. These areas benefit most from models that treat internal psychological states as dynamic processes rather than static classifications. As sensor technology, data availability, and deep learning based latent models continue to advance, DLSMs are expected to play an even more central role in building intelligent systems that understand, predict, and respond to the fluid nature of human behavior.

Benchmark Datasets Used

Benchmark datasets play a crucial role in advancing Dynamic Latent State Models (DLSMs) for behavioral and physiological prediction, as they enable standardized evaluation, reproducibility, and fair comparison across modeling approaches. Over the past decade, several datasets have emerged as foundational resources in the field, particularly those offering multimodal physiological signals and dynamic behavioral markers. These datasets provide diverse annotations ranging from stress labels and emotional states to cognitive workload levels that align well with the latent-state perspective adopted in DLSM research.

One of the most widely used datasets in this domain is the DEAP (Dataset for Emotion Analysis using Physiological Signals), which includes multimodal recordings such as EEG, electrodermal activity (EDA), electromyography (EMG), and respiration signals. DEAP has become a benchmark for emotion recognition tasks and is frequently used in studies developing dynamic latent state models for affect inference[22]. Its rich temporal resolutions make it particularly suitable for DLSMs, which rely on sequential patterns to infer transitions between emotional states. Researchers often leverage DEAP to evaluate dynamic models like Hidden Markov Models, deep Kalman filters, and variational state-space networks in tasks involving arousal and valence prediction.

Another prominent benchmark is the AMIGOS dataset, which focuses on affect, personality, and group-level emotional responses. AMIGOS includes synchronized EEG, EDA, ECG, and video recordings, providing a multimodal environment that supports the development of latent models capable of capturing interpersonal and intrapersonal emotional dynamics[23]. The dataset's design, which includes both individual and group scenarios, makes it particularly valuable for studies investigating complex temporal emotional interactions.

For stress and cognitive workload research, the WESAD (Wearable Stress and Affect Detection) dataset has become a cornerstone. WESAD provides high-quality physiological recordings, including ECG, EDA, respiration, body temperature, and accelerometer data. Its labeling of baseline, stress, and amusement states enables researchers to evaluate how well dynamic models detect transitions in autonomic arousal. Leveraging WESAD, DLSMs have contributed to improving continuous stress detection, especially in real-time applications where physiological dynamics shift gradually rather than abruptly.

Similarly, the MAHNOB-HCI dataset has played a central role in affective computing and human computer interaction research. This dataset includes synchronized EEG, ECG, breathing rate, skin temperature, and facial video recordings[24]. With detailed emotional annotations based on both self-reports and facial behavior coding, MAHNOB-HCI is well-suited for validating dynamic latent models that integrate multimodal data streams. The dataset's high temporal fidelity supports the modeling of fine-grained emotional transitions that are difficult to capture using static classifiers.

In fatigue and cognitive performance research, datasets such as the SEED and SEED-IV datasets have gained prominence. SEED provides multi-session EEG recordings with labeled emotional and cognitive states, enabling longitudinal studies with dynamic models. SEED-IV extends this with four emotion categories and higher-resolution signals. These datasets are frequently used to evaluate deep

recurrent state-space models and neural ODE-based latent models, which aim to capture long-term temporal dependencies across multiple recording sessions[25].

More recent work has also incorporated datasets from naturalistic or real-world contexts, reflecting growing interest in ecological validity. For example, the DREAMER dataset includes EEG and ECG data collected in a realistic emotional induction setting, while AffectiveROAD provides mobile physiological sensor recordings in driving scenarios. These datasets support the development of DLsMs that generalize beyond controlled laboratory environments, an essential step toward real-world deployment.

Collectively, these benchmark datasets have shaped the evolution of dynamic latent state modeling by providing diverse, high-quality physiological and behavioral data. Their widespread use reflects not only their methodological richness but also their alignment with the objectives of DLsMs: understanding how internal states evolve over time. As new datasets emerge with richer multimodal content and real-world conditions, they will continue to drive advancements in dynamic latent modeling and broaden the horizons of behavioral prediction research.

Comparison of the Results of Current Research with Previous Studies

A comparison between the findings of the current mapping study and those of previous research reveals both continuity and important advancements in the understanding and application of Dynamic Latent State Models (DLsMs) within behavioral and physiological prediction tasks. Earlier studies over the past decade primarily focused on evaluating individual dynamic models such as Hidden Markov Models (HMMs), Linear Dynamical Systems (LDS), and classical Kalman filters for predicting discrete or linear physiological and psychological states. These foundational works consistently reported that dynamic models outperform static machine learning approaches when dealing with time-dependent biosignals such as EEG, HRV, and EDA, largely due to their ability to capture temporal patterns and latent internal processes. However, prior work tended to evaluate models in isolated contexts and rarely provided a comprehensive overview of the field's evolution across model types, applications, or datasets.

The current research expands on previous findings by offering a systematic and longitudinal perspective on how DLsMs have evolved over the past ten years[26]. While earlier studies noted the limitations of linear or discrete-state models, the present mapping study quantifies the shift toward deep learning based latent state models, such as Deep Kalman Filters (DKFs), Variational State-Space Models (VSSMs), and Neural ODE latent models. Unlike prior works that highlighted individual model performance, this study demonstrates a clear trend: after 2020, deep latent state models became the dominant methodology, surpassing classical probabilistic approaches not only in accuracy but also in scalability, flexibility, and multimodal integration. This trend was only hinted at in earlier research but is now supported by substantial empirical evidence.

Additionally, previous studies often focused on single applications most commonly emotion recognition or stress detection leading to fragmented insights across domains. The current research consolidates these findings and shows that stress, emotion, and fatigue remain the three dominant application areas, but the scope of DLsM applications has broadened significantly. Unlike earlier work, which treated these domains separately, this study maps how model choices differ across applications, revealing that EEG-based emotion recognition relies heavily on nonlinear state-space models, while HRV and EDA-based stress detection increasingly uses deep generative latent models with uncertainty estimation. This level of cross-domain comparison was largely absent in previous literature.

Another key contribution of the present research is its systematic analysis of dataset usage. Earlier studies tended to highlight the strengths of individual datasets such as DEAP or WESAD but did not examine how dataset selection influences model development across the field[27]. In contrast, this mapping study shows that specific datasets (e.g., DEAP for emotion, WESAD for stress, SEED for cognitive and emotional tasks) have shaped the evolution of DLsMs by providing benchmark environments that encouraged the adoption of more advanced latent models. This dataset-centric perspective offers new insights into how research ecosystems evolve.

Finally, whereas earlier works typically emphasized model performance, the current research provides a broader overview of methodological trends, research density over time, emerging technologies, and gaps in the literature. The study reveals that although deep latent state models have surged in popularity, challenges remain in interpretability, real-world deployment, multimodal synchronization, and generalization across populations issues that earlier research acknowledged but did not systematically categorize.

4. Conclusion

This mapping study provides a comprehensive synthesis of research on Dynamic Latent State Models (DLSMs) in behavioral and physiological prediction over the past decade. The findings show a clear and consistent growth in the adoption of DLSMs, reflecting increasing recognition of the importance of modeling temporal dynamics, hidden states, and probabilistic uncertainty in human behavior and physiological processes. Compared with earlier work that predominantly used static machine-learning models or basic probabilistic approaches, recent studies demonstrate a shift toward more expressive deep latent state models, such as deep Kalman filters and variational state-space models, which offer greater flexibility in handling high-dimensional and nonstationary data. The analysis reveals that EEG, HRV, and EDA remain the most frequently used physiological signals, and that stress, emotion, and fatigue prediction are the dominant application domains. Benchmark datasets such as DEAP, WESAD, and DREAMER play a central role in shaping methodological development, though the field still suffers from limited ecological diversity in available data. The comparison with prior research indicates that while earlier studies emphasized interpretability and structured probabilistic modeling, current research prioritizes scalability, multimodality, and deep generative representation learning. Despite the progress, several challenges persist, including interpretability limitations in deep latent models, inconsistent evaluation protocols, and the scarcity of large, real-world longitudinal datasets. These gaps highlight opportunities for future research, particularly in developing hybrid interpretable deep modeling frameworks, improving transparency in latent-state inference, and expanding accessible datasets to support more robust benchmarking. Overall, this study contributes a structured overview of theoretical foundations, methodological trends, and practical applications of DLSMs, offering a consolidated reference for researchers and practitioners. The insights generated serve as a roadmap for advancing dynamic latent-state modeling toward more accurate, interpretable, and generalizable behavioral and physiological prediction systems.

References

- [1] J. L. Andreassi, *Psychophysiology: Human behavior and physiological response*. Psychology press, 2010.
- [2] P. Rajendra and V. Brahmajirao, "Modeling of dynamical systems through deep learning," *Biophys. Rev.*, vol. 12, no. 6, pp. 1311–1320, 2020.
- [3] F. Chen *et al.*, *Robust multimodal cognitive load measurement*. Springer, 2016.
- [4] A. Klushyn, R. Kurlle, M. Soelch, B. Cseke, and P. van der Smagt, "Latent matters: Learning deep state-space models," *Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 10234–10245, 2021.
- [5] C. Chen, "Science mapping: a systematic review of the literature," *J. data Inf. Sci.*, vol. 2, no. 2, 2017.
- [6] S. M. Siddiqi, "Latent Variable and Predictive Models of Dynamical Systems," 2007.
- [7] C. Esteban, O. Staeck, S. Baier, Y. Yang, and V. Tresp, "Predicting clinical events by combining static and dynamic information using recurrent neural networks," in *2016 IEEE international conference on healthcare informatics (ICHI)*, IEEE, 2016, pp. 93–101.
- [8] S. Lüdtke, M. Schröder, F. Krüger, S. Bader, and T. Kirste, "State-space abstractions for probabilistic inference: a systematic review," *J. Artif. Intell. Res.*, vol. 63, pp. 789–848, 2018.
- [9] I. M. Miale-Lye, S. Hempel, R. Shanman, and P. G. Shekelle, "What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products," *Syst. Rev.*, vol. 5, no. 1, p. 28, 2016.
- [10] D. B. Rice, L. A. Kloda, I. Shrier, and B. D. Thombs, "Reporting completeness and transparency of meta-analyses of depression screening tool accuracy: a comparison of meta-analyses published before and after the PRISMA statement," *J. Psychosom. Res.*, vol. 87, pp. 57–69, 2016.

- [11] D. Papadopoulos, N. Papadakis, and A. Litke, "A methodology for open information extraction and representation from large scientific corpora: the CORON-19 data exploration use case," *Appl. Sci.*, vol. 10, no. 16, p. 5630, 2020.
- [12] A. Sattar Chaudhry, "Assessment of taxonomy building tools," *Electron. Libr.*, vol. 28, no. 6, pp. 769–788, 2010.
- [13] H. Hermessi, O. Mourali, and E. Zagrouba, "Deep feature learning for soft tissue sarcoma classification in MR images via transfer learning," *Expert Syst. Appl.*, vol. 120, pp. 116–127, 2019.
- [14] E. Sükei, A. Norbury, M. M. Perez-Rodriguez, P. M. Olmos, and A. Artés, "Predicting emotional states using behavioral markers derived from passively sensed data: data-driven machine learning approach," *JMIR mHealth uHealth*, vol. 9, no. 3, p. e24465, 2021.
- [15] C. Chen, C. X. Lu, B. Wang, N. Trigoni, and A. Markham, "DynaNet: Neural Kalman dynamical model for motion estimation and prediction," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 32, no. 12, pp. 5479–5491, 2021.
- [16] G. Valenza and E. P. Scilingo, "Autonomic nervous system dynamics for mood and emotional-state recognition: Significant advances in data acquisition, signal processing and classification," 2013.
- [17] J. Zhu, L. Ji, and C. Liu, "Heart rate variability monitoring for emotion and disorders of emotion," *Physiol. Meas.*, vol. 40, no. 6, p. 64004, 2019.
- [18] H. F. Posada-Quintero and K. H. Chon, "Innovations in electrodermal activity data collection and signal processing: A systematic review," *Sensors*, vol. 20, no. 2, p. 479, 2020.
- [19] M. W. Loftis and P. B. Mortensen, "A dynamic linear modelling approach to public policy change," *J. Public Policy*, vol. 38, no. 4, pp. 553–579, 2018.
- [20] X. Zhang, W. Li, X. Chen, and S. Lu, "Moodexplorer: Towards compound emotion detection via smartphone sensing," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–30, 2018.
- [21] S. Kerick, J. Metcalfe, T. Feng, A. Ries, and K. McDowell, "Review of fatigue management technologies for enhanced military vehicle safety and performance," 2013.
- [22] R. Smith, T. Parr, and K. J. Friston, "Simulating emotions: An active inference model of emotional state inference and emotion concept learning," *Front. Psychol.*, vol. 10, p. 2844, 2019.
- [23] S. Siddharth, *Utilizing Multi-modal Bio-sensing Toward Affective Computing in Real-world Scenarios*. University of California, San Diego, 2020.
- [24] K. Sharma, C. Castellini, E. L. Van Den Broek, A. Albu-Schaeffer, and F. Schwenker, "A dataset of continuous affect annotations and physiological signals for emotion analysis," *Sci. data*, vol. 6, no. 1, p. 196, 2019.
- [25] L. Allen, J. Jiang, and J. Lubuma, "ICMA VII: The Seventh International Conference on Mathematical Modeling and Analysis of Populations in Biological Systems," 2019.
- [26] L. Ortega and G. Iberri-Shea, "Longitudinal research in second language acquisition: Recent trends and future directions," *Annu. Rev. Appl. Linguist.*, vol. 25, pp. 26–45, 2005.
- [27] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing wesad, a multimodal dataset for wearable stress and affect detection," in *Proceedings of the 20th ACM international conference on multimodal interaction*, 2018, pp. 400–408.