

Paper	Year	Dataset	Objectives	Techniques
Héron: Taming Tail Latencies in Key-Value Stores Under Heterogeneous Workloads	2018	Wikipedia	Workload characterization	Optimization (bloom filters)
A Long-term Cloud Workload Prediction Framework for Reserved Resource Allocation	2022	Wikipedia	Workload prediction	Deep learning (transformer-based)
Workload prediction using ARIMA model and its impact on cloud applications' QoS	2014	Wikipedia	Workload prediction	Time-series models (ARIMA)
An adaptive workload prediction strategy for non-Gaussian cloud service using ARMA model with higher order statistics	2018	Wikipedia	Workload prediction	Time-series models (ARMA), Statistical analysis
CloudInsight: Utilizing a Council of Experts to Predict Future Cloud Application Workloads	2018	Wikipedia	Workload prediction	Time-series models (ARIMA, ARMA, AR, BDES, WMA), Classical machine learning (LR, SVM)
CANFIS: A Chaos Adaptive Neural Fuzzy Inference System for Workload Prediction in the Cloud	2022	Wikipedia, Nasa	Workload prediction	Filtering and Signal Processing (Savitzky-Golay (SG) filter), time series analysis, Deep learning (Adaptive Neural Fuzzy Inference System (ANFIS))
Wikipedia workload analysis for decentralized hosting	2009	Wikipedia	Workload characterization; Workload classification	Time series analysis
Chameleon: A Hybrid, Proactive Auto-Scaling Mechanism on a Level-Playing Field	2018	Wikipedia, WorldCup98, RetailRocket	Workload prediction; Auto-scaling	Time-series models (sARIMA, TBATS)
Higher Order Statistics Based Method For Workload Prediction In The Cloud Using ARMA Model	2018	Wikipedia	Workload prediction	Statistical analysis, Time-series models (Autoregressive Moving Average (ARMA))
Online Workload Burst Detection for Efficient Predictive Autoscaling of Applications	2020	Wikipedia, WorldCup98, Nasa, Calgary, Clarknet	Workload prediction; Auto-scaling	Classical machine learning (Linear Regression (LR), Polynomial Regression (PR), Elastic Net (EN) Regression, Ridge Regression, Lasso Regression, XGBoost (XGB) Regression, Random Decision Forests (RDF) Regression, and Decision Tree Regression (DTR)), Time-series models (Sample Entropy (SE), Normalized Entropy (NE)), Filtering and Signal Processing (FFT);
CloudBruno: A Low-Overhead Online Workload Prediction Framework for Cloud Computing	2022	Wikipedia	Workload prediction	Deep learning (LSTM), Classical machine learning (SVM)
A Self-Optimized Generic Workload Prediction Framework for Cloud Computing	2020	Wikipedia	Workload prediction	Deep learning (LSTM), Optimization (Bayesian Optimization)
Self-adaptive workload classification and forecasting for proactive resource provisioning	2013	Wikipedia	Workload prediction; Workload classification	Time-series models (Naive forecast, MA, ARIMA, Croston's Method, TBATS, sARIMA, Simple Exponential Smoothing (SES), Extended Exponential Smoothing (ETS)) Filtering and Signal Processing (Cubic Smoothing Splines (CS));
PEAS: A Performance Evaluation Framework for Auto-Scaling Strategies in Cloud Applications	2016	Wikipedia	Benchmarking (Auto-scaling)	Control Theory, Optimization (Stochastic Programming)
Workload Classification for Efficient Auto-Scaling of Cloud Resources	2013	Wikipedia, WorldCup98	Workload classification	Time series models (Sample Entropy), Classical machine learning (KNN)
A Fuzzy Approach Based on Heterogeneous Metrics for Scaling Out Public Clouds	2017	Wikipedia, WorldCup98	Auto-scaling	Fuzzy logic, Control theory (Proportional Integrative Derivative (PID))
A new temporal locality-based workload prediction approach for SaaS services in a cloud environment	2022	Wikipedia	Workload prediction	Others (Temporal locality)
Burst-Aware Predictive Autoscaling for Containerized Microservices	2020	Wikipedia, WorldCup98, Nasa, Calgary, Clarknet	Workload prediction; Auto-scaling	Classical machine learning (Linear Regression(LR), Polynomial Regression(PR), Elastic Net (EN) regression, XGBoost (XGB) regression, Random Decision Forests (RDF) regression, Decision Tree Regression (DTR), Support Vector regression (SVR)), Deep learning (Multi-layer Perceptron regression (MLP)), Filtering and Signal Processing (FFT);
Virtual Machine Provisioning Based on Analytical Performance and QoS in Cloud Computing Environments	2011	Wikipedia	Resource provisioning	Queueing theory
Learning-Based Power Prediction for Data Centre Operations via Deep Neural Networks	2016	WorldCup98, Clarknet	Workload prediction	Deep learning (Auto Encoder (AE), Recursive Auto Encoder (RAE))
A Workload Prediction Approach Using Models Stacking Based on Recurrent Neural Network and Autoencoder	2016	WorldCup98, Nasa	Workload prediction	Deep learning (GRU, AE)
Efficient Autoscaling in the Cloud using Predictive Models for Workload Forecasting	2011	WorldCup98	Workload prediction, Resource provisioning	Time-series models (ARMA); Others (Graph-based), Control theory (look-ahead Controller)
A Workload Characterization Study of the 1998 World Cup Web Site	2000	WorldCup98	Workload characterization	Time series analysis

PRESS: PRedictive Elastic ReSource Scaling for cloud systems	2010	WorldCup98, Clarknet	Workload prediction; Auto-scaling	Markov model (Markov chain model), Filtering and Signal processing (FFT);
Combining time series prediction models using genetic algorithm to autoscaling Web applications hosted in the cloud infrastructure	2016	WorldCup98, Nasa, Clarknet	Workload prediction; Auto-scaling	Time-series models (Naive forecast, AR, MA, ARMA, Holt-winters, ARIMA, Simple exponential smoothing (SES), Extended exponential smoothing (ETS)); Optimization (Genetic algorithm)
K-AGRUEd: A Container Autoscaling Technique for Cloud-based Web Applications in Kubernetes Using Attention-based GRU Encoder-Decoder	2022	WorldCup98, Nasa	Workload prediction; Auto-scaling	Deep learning (transformer-based, GRU, BiGRU);
A hybrid wavelet decomposer and GMDH-ELM ensemble model for Network function virtualization workload forecasting in cloud computing	2020	WorldCup98	Workload prediction	Filtering and Signal Processing (discrete wavelet transform DWT), Deep learning (ANN)
Machine learning-based auto-scaling for containerized applications	2020	WorldCup98	Workload prediction; Auto-scaling	Deep learning (LSTM)
Injecting Realistic Burstiness to a Traditional Client-Server Benchmark	2009	WorldCup98	Benchmarking (Burst injection)	Queueing theory (Markovian Arrival Processes (MAP))
Analytic Modeling of Load Balancing Policies for Tasks with Heavy-tailed Distributions	2000	WorldCup98	Load balancing	Queueing theory (Markov chain model)
Flexible and Efficient Decision-Making for Proactive Latency-Aware Self-Adaptation	2018	WorldCup98, Clarknet	self-healing	Optimization (Stochastic Programming), Markov model (Markov decision processes (MDP))
TeaStore: A Micro-Service Reference Application for Benchmarking, Modeling and Resource Management Research	2018	WorldCup98	Benchmarking (system)	Others (System design)
A QoS-Based Reactive Auto Scaler for Cloud Environment	2017	WorldCup98	Auto-scaling	Queueing theory
Efficient Decision-Making under Uncertainty for Proactive Self-Adaptation	2016	WorldCup98, Clarknet	self-healing	Optimization (Stochastic Programming), Markov model (Markov decision processes (MDP))
X-RAY: A Non-Invasive Exclusive Caching Mechanism for RAIDs	2004	WorldCup98, Nasa	Caching optimization	Others (Cache Management Policies)
Research on scheduling algorithms in Web cluster servers	2003	WorldCup98, Nasa, Calgary, Clarknet	Load balancing	Time series analysis
Deep Learning-Based Autoscaling Using Bidirectional Long Short-Term Memory for Kubernetes	2021	WorldCup98, Nasa	Workload prediction; Auto-scaling	Deep learning (Bi-LSTM)
A Web Application Load Prediction Model Using Recurrent Neural Network in Cloud	2020	WorldCup98	Workload prediction	Deep learning (Bi-LSTM)
Resource provisioning using workload clustering in cloud computing environment: a hybrid approach	2021	WorldCup98, Nasa	Resource provisioning; Workload Classification	Classical machine learning (decision tree); Optimization (Imperialist Competition Algorithm (ICA)), Classical machine learning (KMeans)
EquiLoad: a load balancing policy for clustered web servers	2001	WorldCup98	Load Balancing	Queueing theory
Optimized Hierarchical Tree Deep Convolutional Neural Network of a Tree-Based Workload Prediction Scheme for Enhancing Power Efficiency in Cloud Computing	2023	Nasa, Saskatchewan	Workload prediction	Deep learning (CNN), Classical machine learning (Kernel method), Optimization (sheep flock optimization (SFO))
A Quantum Approach Towards the Adaptive Prediction of Cloud Workloads	2021	Nasa, Saskatchewan	Workload prediction	Deep learning (Quantum neural network)
BiPhase adaptive learning-based neural network model for cloud datacenter workload forecasting	2020	Nasa, Saskatchewan	Workload prediction	Deep learning (ANN), Optimization (Different evolution algorithm)
Workload prediction in cloud using artificial neural network and adaptive differential evolution	2018	Nasa, Saskatchewan	Workload prediction	Deep learning (ANN), Optimization (Differential evolution algorithm)
Self directed learning based workload forecasting model for cloud resource management	2021	Nasa, Saskatchewan, Calgary	Workload prediction	Deep learning (MLP), Optimization (Blackhole algorithm)
Auto-adaptive learning-based workload forecasting in dynamic cloud environment	2022	Nasa, Saskatchewan	Workload prediction	Deep learning (ANN), Optimization (Different evolution algorithm)
Long short term memory recurrent neural network (LSTM-RNN) based workload forecasting model for cloud datacenters	2018	Nasa, Saskatchewan, Calgary	Workload prediction	Deep learning (LSTM)
Prediction of Cloud Data Center Networks Loads Using Stochastic and Neural Models	2011	Nasa, EPA	Workload prediction	Deep learning (MLP), Time-series models (autoregressive (AR)), Filtering and Signal Processing (Weiner Filter)
Dynamic Resource Scaling in Cloud Using Neural Network and Black hole Algorithm	2016	Nasa, Saskatchewan, Calgary	Workload prediction	Deep learning (ANN), Optimization (Black hole algorithm)
FAHP approach for autonomic resource provisioning of multitier applications in cloud computing environments	2018	Nasa, Clarknet	Resource provisioning; Workload prediction	Fuzzy logic, Control theory (MAPE-K control loop); Classical machine learning (Support vector regression (SVR))

Web Server Workload Characterization: The Search for Invariants	1996	Nasa, Saskatchewan, Calgary, Clarknet	Workload characterization	Time series analysis
Performance evaluation of metaheuristics algorithms for workload prediction in cloud environment	2021	Nasa, Saskatchewan, Calgary	Benchmarking (Workload prediction)	Optimization (Genetic algorithm, Differential evolution, Blackhole algorithm, others), Statistical analysis
Performance Assessment of Time Series Forecasting Models for Cloud Datacenter Networks' Workload Prediction	2021	Nasa, Saskatchewan, Calgary	Benchmarking (Workload prediction)	Time-series models (AR, MA, ARMA, ARIMA, Auto ARIMA, Exponential Smoothing (ES))
A cloud computing load prediction hybrid model with adaptive weight strategy	2023	Nasa, Saskatchewan	Workload prediction	Filtering and Signal Processing (empirical mode decomposition (EMD)), Deep learning (LSTM, convolution neural network (CNN), transformer-based)
RHAS: robust hybrid auto-scaling for web applications in cloud computing	2021	Nasa, Clarknet	Workload prediction; Auto-scaling	Time-series models (ARIMA), Classical machine learning (SVR), time-series analysis;
A Preliminary Study of Machine Learning Workload Prediction Techniques for Cloud Applications	2019	Nasa	Workload prediction	Time-series models (ARIMA), Deep learning (MLP, GRU)
Application of an Enhanced Self-adapting Differential Evolution Algorithm to Workload Prediction in Cloud Computing	2019	Nasa, Saskatchewan	Workload prediction	Deep learning (ANN), Optimization (Differential evolution algorithm)
Cloud datacenter workload estimation using error preventive time series forecasting models	2020	Nasa, Saskatchewan, Calgary	Workload prediction	Statistical analysis
TASM: technocrat ARIMA and SVR model for workload prediction of web applications in cloud	2019	Nasa, Clarknet	Workload prediction	Classical machine learning (linear regression, SVR), Time-series models (ARIMA), time-series analysis, Optimization (Differential evolution algorithm)
Lognormal and Pareto distributions in the Internet	2005	Nasa, Saskatchewan, Calgary, Clarknet, Boston	Workload characterization	Statistical analysis
Workload prediction for enhancing power efficiency of cloud data centers using optimized self-attention-based progressive generative adversarial network	2024	Nasa, Saskatchewan	Workload prediction	Markov model (Markov random field (MRF)), Deep learning (Generative adversarial network (GAN), transformer-based), Optimization (others)
COSCO2: AI-augmented evolutionary algorithm based workload prediction framework for sustainable cloud data centers	2023	Nasa, Saskatchewan	Workload prediction	Deep learning (CNN), Classical machine learning (Kernel method), Optimization (sheep flock optimization (SFO))
Autonomic task scheduling algorithm for dynamic workloads through a load balancing technique for the cloud-computing environment	2021	Nasa, Calgary	Load Balancing	Deep learning (ANFIS), Classical machine learning (Naive Bayesian Classification)
Toward on-Line Predictive Models for Forecasting Workload in Clouds	2018	Clarknet	Workload prediction	Classical machine learning (SGDRegressor)
An Online Control Framework for Designing Self-Optimizing Computing Systems: Application to Power Management	2004	Clarknet	Auto-scaling	Control theory (look-ahead Controller)
Conserving Disk Energy in Network Servers	2003	Clarknet	Benchmarking (Power efficiency)	Others (Disk Management Strategies), Time-series models (Exponentially Weighted Moving Average (EWMA))
Category-based User Interaction with Online User-Generated Videos: Workload Characterization	2014	Youtube	Workload characterization	Time series analysis
Characteristics of YouTube network traffic at a campus network – Measurements, models, and implications	2009	Youtube	Workload characterization	Time series analysis
Efficient resource provisioning for elastic Cloud services based on machine learning techniques	2019	Madrid	Resource provisioning	Classical machine learning (SVR, Kernel method), Queuing theory
Cushion: A proactive resource provisioning method to mitigate SLO violations for containerized microservices	2022	WorldCup98, Wikipedia, Nasa, Clarknet	Resource provisioning; Auto-scaling	Deep learning (LSTM), Classical machine learning (XGBoost)
Predictive Autoscaling of Microservices Hosted in Fog Microdata Center	2020	Wikipedia, WorldCup98, Clarknet	Workload prediction; Auto-scaling	Classical machine learning (Elastic net (EN), DTR)
Auto-Scaling Containerized Applications in Geo-Distributed Clouds	2023	Wikipedia, Nasa	Workload prediction; Auto-scaling	Deep learning (Deep reinforcement learning (DRL))
Metrics for Mass-Count Disparity	2006	SDSC, WorldCup98	Workload characterization	Time series analysis
Web workload characterization: Ten years later	2005	Calgary, Saskatchewan	Workload characterization	Time series analysis
Watch Global Cache Local: YouTube Network Traces at a Campus Network - Measurements and Implications	2008	Youtube	Workload characterization	Time series analysis
Characteristics of WWW Client-based Traces	1995	Boston	Workload characterization	Time series analysis
Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes	1997	Boston	Workload characterization	Time series analysis
Changes in Web client access patterns: Characteristics and caching implications	1999	Boston	Workload characterization; Caching optimization	Time series analysis