Journal of Grid Computing (2024) 22:65
https://doi.org/10.1007/s10723-024-09780-4

RESEARCH

Assessing the Complexity of Cloud Pricing Policies: A Comparative Market Analysis

Vasiliki Liagkou · George Fragiadakis · Evangelia Filiopoulou · Christos Michalakelis · Anargyros Tsadimas · Mara Nikolaidou

Received: 25 March 2024 / Accepted: 2 September 2024

Abstract Cloud computing has gained popularity at a breakneck pace over the last few years. It has revolutionized the way businesses operate by providing a flexible and scalable infrastructure for their computing needs. Cloud providers offer a range of services with a variety of pricing schemes. Cloud pricing schemes are based on functional factors like CPU, RAM, and storage, combined with different payment options, such as pay-per-use, subscription-based, and non-functional aspects, such as scalability and availability. While cloud pricing can be complicated, it is critical for businesses to thoroughly assess and compare pricing policies along with technical requirements to ensure they design an investment strategy. This paper evaluates current pricing strategies for IaaS, CaaS, and PaaS cloud services and also focuses on the three leading cloud providers, Amazon, Microsoft, and Google. To compare pricing policies between different services and providers, a hedonic price index is constructed for each service type based on data collected in 2022. Using the hedonic price index, a comparative analysis between them becomes feasible. The results revealed that providers follow the very same pricing pattern for IaaS and CaaS, with CPU being the main driver of cloud pricing schemes, whereas PaaS pricing fluctuates among cloud providers.

Keywords Pricing policy models · Hedonic price index · IaaS · CaaS · PaaS · Cloud market

George Fragiadakis, Evangelia Filiopoulou, Christos Michalakelis, Anargyros Tsadimas and Mara Nikolaidou contributed equally to this work.

V. Liagkou · G. Fragiadakis (✉) · E. Filiopoulou · C. Michalakelis · A. Tsadimas · M. Nikolaidou
Department of Informatics & Telematics, Harokopio University of Athens, Omirou 9, Tavros 17778, Attika, Greece
e-mail: gfragi@hua.gr

V. Liagkou
e-mail: vliagkou@hua.gr

E. Filiopoulou
e-mail: evangel@hua.gr

C. Michalakelis
e-mail: michalak@hua.gr

A. Tsadimas
e-mail: tsadimas@hua.gr

M. Nikolaidou
e-mail: mara@hua.gr

1 Introduction

Cloud computing services have expanded significantly in response to the ever-evolving demands of technological advancements, offering sophisticated solutions that cater to a wide range of needs and industries. Cloud providers are continuously striving to stay at the forefront by rapidly expanding their service offerings, each bringing a unique set of features and innovations to the table. These range from foundational services like Infrastructure as a Service (IaaS) and Platform as a Ser-

Published online: 12 September 2024

vice (PaaS) to more specialized options such as Container as a Service (CaaS), offering diverse approaches to infrastructure and application management.

Each of these cloud services has its own pricing policies, incorporating different factors affecting the price of each service. They are provided in the form of prescribed bundles and combine multiple functional and non-functional features. This makes it challenging for customers to select the right services and discern what they are paying for. Surveys have consistently shown that this lack of clarity is a significant pain point. The primary challenges for customers, as shown in [1], include how pricing is accounted for, measured in the cloud, and spread across the various service layers. Anodot State of Cloud Cost Report revealed that 50% of cloud-based businesses face challenges in controlling cloud costs, with the complexity of cloud pricing and a lack of visibility driving cloud costs being the main issues [2]. Similarly, according to [3], only 30% of the organizations surveyed knew exactly where their cloud budget was going. Last but not least, 96% of IT decision makers said that cost control is one of the significant challenges faced when implementing their cloud strategy [4].

Toward this direction, this paper aims to assess the contribution of prominent cloud service features to the service bundle pricing policy and provide a comparative analysis of pricing policies across IaaS, CaaS, and PaaS cloud service types and the three leading providers. To enable the comparison, a common function to describe pricing policies is needed. Based on the authors' previous works [5–7], we suggest using a pricing model based on hedonic principles and constructing price indices for IaaS, CaaS, and PaaS services across major cloud providers. This hedonic price index approach is particularly appealing as it allows for the comparison of common cloud features among different services.

This study offers a direct contribution to cloud service pricing by providing a detailed comparison of IaaS, CaaS, and PaaS pricing policies as well as an analysis of pricing strategy variations amongst providers. This comparison is crucial for decision-makers, as it sheds light on the pricing policies of the three service types required for application development. Additionally, it enables users to assess the value proposition of each cloud service type more comprehensively, considering not only the absolute pricing but also the relative cost-effectiveness of individual features. Understanding the disparity in pricing contributions of features between cloud service types can guide users in selecting the most cost-effective service type and vendor for their particular workload requirements. This knowledge empowers them to facilitate strategic planning and budget optimization by highlighting areas where potential cost savings or efficiency improvements might be obtained.

The paper is organized as follows: Section 2 summarizes related work. The features of cloud services that participate in pricing schemes are covered in Section 3. The way hedonic indices are developed to predict cloud pricing strategies is briefly discussed in Section 4, while results are examined in Section 5. A discussion based on a comparison analysis across service categories and cloud providers is provided in Section 6. Conclusion and future work sit in Section 7.

2 Motivation and Related Work

Before introducing the hedonic pricing analysis, it is necessary to discuss the cloud computing environment. Following this discussion, related work will be reviewed, focusing on academic studies concerning pricing models.

2.1 Cloud Computing Landscape

Understanding the pricing policies of popular cloud providers necessitates a basic understanding of the cloud market and the wide range of products available. This background is critical to determining how these products are priced and how various pricing strategies affect customers. Following this examination of the cloud computing landscape, we will look at relevant research into various areas of cloud service pricing, such as economic models, consumer behavior, and the technology breakthroughs driving these changes.

Cloud computing, defined by the Institute of Standards and Technology (NIST) [8] and the ITU-T Recommendation Y.3500 [9] as the *on-demand availability of computer system resources*, has revolutionized how businesses access and utilize technology. This technology has significantly transformed the way businesses utilize and access technology. The primary entities in this domain are cloud providers, which provide the services, and cloud customers, who use the services [10].

In this constantly evolving environment, cloud providers provide many service delivery models, each containing a set of features tailored to meet distinct user requirements. According to the Institute of Standards and Technology (NIST) [8], there are three basic cloud service types, which include Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). While the latest targets application users, the first two are used by application developers when designing a cloud solution.

- Infrastructure as a Service (IaaS): Provides the fundamental building blocks of computing—virtual machines, storage, and networking—enabling consumers to build customized environments while retaining control over operating systems and applications.
- Platform as a Service (PaaS): simplifies application development by offering a comprehensive platform encompassing infrastructure, operating systems, middleware, and development tools. This abstraction empowers developers to focus solely on building and deploying applications.

According to [11], applications are increasingly being developed to be cloud-ready, adhering to the patterns established by the Cloud Native Computing Foundation (CNCF). Furthermore, cloud applications typically adopt a microservices architectural pattern, moving away from the traditional monolithic approach. Virtualization alone (e.g. IaaS) is insufficient to build native cloud applications [12]. The implementation of DevOps principles further facilitates this transformation. Monolithic applications are typically deployed on cloud VMs, whereas microservice architectures are now delivered as containers and run on container orchestrators like Kubernetes. This led to the increase use of a new subtype, Container as a Service (CaaS).

Since the National Institute of Standards and Technologies (NIST) has not been updated the cloud computing definitions since 2011 [13] and cloud technology has advanced significantly, delivery methodologies and commercial offerings have evolved.

In ITU-T Y.3500 published in 2014, the acronym CaaS refers to Communication as a Service, a cloud service category not widely adopted in cloud service market. Container as a Service, abbreviated to CaaS by cloud providers' industry, typically refers to a cloud service model that allows users to manage and deploy containerized applications and services. It includes features like orchestration, scaling, and deployment of containers. In every day practice, taking into account the number of products available by cloud providers, in our days the abbreviation CaaS commonly referred to "Container as a Service" rather than "Communication as a Service" , though both service categories are valid.

As we acknowledge the importance of aligning our work with existing recommendations, we would like to note that "Container as a Service" category may be classified as a specialized subcategory of ITU-T Y.3500 "Compute as a Service" category, abbreviated to CompassS, referring the provision and use of processing resources needed to deploy and run software.

- Container as a Service (CaaS): leverages containerization technology for efficient application deployment and management. Containers offer a lightweight, portable, and scalable alternative to virtual machines, optimizing resource utilization. Although CaaS can be considered a subset of PaaS due to its specific focus on container management, it remains a distinct service from the consumer's perspective as it addresses unique needs and use cases. [14] highlights how higher-level services like Containers-as-a-Service (CaaS) build upon fundamental layers like Infrastructure-as-a-Service (IaaS) and offers insights into the tiered nature of cloud service architectures.

Developers can build an application utilizing either IaaS, PaaS or CaaS products. Each discrete product, called service bundle by cloud service providers, consists of a set of specific functional and non-functional features and billing options and comes at a specific price. All ad-hear to the same pricing policy in all providers. The number of features belonging a bundle may be numerous and analytically presented by providers, but the user of the service bundle has no knowledge of the way each feature affects the bundle's fixed price. It comes as a black box.

Understanding pricing policies is a powerful strategic tool [15]. A transparent pricing strategy not only builds trust but also positions your business as reliable and straightforward. In this context, this study underlines the importance of comprehending pricing policies for users of cloud services. It highlights the main factors that impact cloud price by simplifying the intricate pricing dynamics of cloud products. It models cloud pricing policy, constructs a hedonic price index, and analyzes the many elements of cloud service bundles in order to identify the characteristics that have the most impact on their prices [16]. A hedonic pricing index facilitates the comprehension of how different characteristics, including both functional features (such as CPU, RAM, and storage) and non-functional attributes (such as scalability, availability, and geographic location), impact the total price. This approach will enable us to determine the attributes that have the most price sensitivity and how alterations in these attributes could potentially result in price fluctuations [17].

2.2 Related Work

The pricing and cost of cloud services are key concerns for many scholars. Pricing a service depends heavily on the cost of production. Therefore, researchers investigate the cost of cloud services as an initial step in understanding pricing. For example, in [18] authors performed an extensive analysis of energy consumption and computational requirements, considering factors such as CPU cycles, memory usage, and storage I/O operations to derive the final pricing values. This comprehensive approach ensures that the cost model reflects the actual resource usage and efficiency of the cloud services. Additionally, in [19] authors conducted a comprehensive assessment of various computational aspects including energy efficiency, resource utilization, and performance metrics to establish the pricing structure. By evaluating these factors, the study provides a detailed understanding of how different elements contribute to the overall cost of cloud services.

In our work we focus on pricing policy exploration and try to explose the importance cloud providers assign to cloud service features when building their pricing policy. We consider this a specific aspect in cloud pricing research that may help cloud solution designers when leverage computational aspects against price in order to build efficient cloud systems. To explore pricing policies we model them using hedonic models. Furthermore, we target to compare policies adopted by different providers. Thus in the following we focus our analysis in a) studies that use the hedonic model to explore pricing policies and b) studies that examine services and prices across providers, as there is no study in the literature focusing on comparing pricing policies across providers.

2.2.1 Hedonic Model: Exploring Pricing Policies

A hedonic index is any price index that uses information from hedonic regression, which describes how product price could be explained by the product's characteristics [16]. It has been widely adopted in several fields, such as in the telecommunications industry [20,21] and in the smartphone market [22].

Moreover, the hedonic method has also been used in cloud environments. In [23], authors proposed two pricing comparison methods. Initially, hedonic pricing was proposed, and then a new method called Pricing Plan Comparison (PriCo) was introduced. The two methods were applied to IaaS services derived from popular providers. Based on the results, Google satisfied users that needed very low CPU and high memory requirements, whereas users that needed high CPU and low memory requirements preferred Terremark. In addition, the results revealed that Google's pricing plan appeared to be a threat to competitors providers.

The authors in [24] described the IaaS services, including intrinsic and extrinsic features. Also, time-dummy variables were integrated into the model, analyzing historical data for Amazon Web Services (AWS) cloud instance prices between 2008 and 2017.

The authors of the present work constructed hedonic price indices based on IaaS services in [5,6]. The IaaS services were initially defined by functional features, such as CPU, RAM, storage, and operating system, in [5], and then non-functional features were also included in [6], highlighting the importance of QoS features in the shaping of the final price. In addition, a price index based on Container-as-a-Service was constructed in [7]. CaaS services were described by functional and non-functional features, and their contribution to the final price was indicated.

2.2.2 Cloud Provider Comparison: Evaluating Pricing Policies

In [25], authors compared the prices of three cloud service providers — Amazon Web Services, Google, and Microsoft Azure - in terms of service availability, data security, operating system, Windows support, free trial, and locations. The comparison was based on IaaS services, and the authors used theoretical literature and an empirical approach. The literature survey was based on the analysis of books, journals, documentation, online studies, etc.

In addition, a comparative pricing analysis between popular cloud providers was introduced in [26]. The authors compared AWS, Azure, Rackspace, HP Cloud, and IBM providers. The IaaS products were categorized into small-scale computation, medium-scale computation, and large-scale computation. For small and medium-sized computing models, AWS offers the most competitive pricing options. However, when it comes to large-scale computing, Azure holds the advantage. In all scenarios, RackSpace tends to be the most expensive option. Finally, based on performance, AWS offers the most cost-effective solutions.

In [27], authors also presented service availability and price comparative analysis of cloud providers (Azure, Amazon, and Oracle). The comparison was conducted across IaaS services using a broad spectrum of criteria, encompassing factors such as the range of services and tools provided by each provider, the platforms with which they were compatible, the languages they accommodated, their security and scalability measures, and the capacity for cloud-based data storage. The authors concluded that there is no way to determine which provider could be considered the optimal one. Amazon had a higher SLA (service level agreement) value than other providers, and Oracle seemed to be the most advantageous choice based on price.

2.3 Contribution

Table 1 summarizes existing work on price policy exploration for cloud services. The presentation of existing work is two-fold.

Firstly, it includes efforts to model pricing policies using price indices based on a diverse number of functional(e.g CPU) and non-functional features (e.g Autoscaling) and various datasets, derived from popular providers. It reveals that the hedonic method has been utilized in IaaS by several scholars, while in CaaS type services only the authors of the current work have examined the pricing policy of the specific service [7].

Secondly, it presents works that explore pricing policies by conducting comparisons between cloud providers' products in terms of price. The related studies either constructed a price index in a specific cloud service or made price comparisons among specific cloud bundles.

Finally, Table 1 includes the current work and its aim to make its contribution clearer to the reader.

The table, besides the method used and the cloud service type that was the focus of each study, also includes information of the dataset used to conduct it, as the type and the number of cloud service features used and the number of different providers and bundles included.

The present study aims to achieve the following contributions:

- Comprehensive comparison of pricing policies: While previous studies have compared cloud prices, none of them compared the pricing policies. They focused on specific and limited bundle comparisons, offering particular findings. Therefore, through a price index, we establish a baseline model capturing the intrinsic value of features within a single cloud provider's pricing structure and then extend this analysis to other major cloud platforms. This allows us to systematically compare and contrast pricing policies across the different cloud services and among various cloud providers, revealing potential patterns, disparities, and strategic considerations.
- Model PaaS pricing policy: This study pioneers the development of a hedonic function for PaaS cloud services, marking a novel contribution to the field.

The findings from this investigation have the potential to significantly inform the design and development of a pricing-oriented decision-making tool. Such a tool would facilitate the optimal selection of cloud services, catering to the specific needs and financial considerations of users. Thus, this study not only contributes to the academic discourse by expanding the application of hedonic models in the cloud computing context but also offers practical implications by aiding stakeholders in making informed decisions regarding cloud service selection.

3 IaaS, CaaS and PaaS Services

Cloud providers offer each service type as a bundle of characteristics at a specific price. Each bundle is a unique combination of functional and non-functional features, which collectively determine the price. Figure 1 illustrates the features of each service type, IaaS, CaaS, and PaaS, that are included and can be adjusted within the bundle pricing. The fixed bundle pricing can vary each time, depending on the specific combination of features selected. This means that

Table 1 Summary of various authors' work with respect to the application of the hedonic model and cloud service comparison

Work	Method	Cloud service type	Feature type	Features (#)	Bundles (#)	Providers (#)
El Kihal et al. (2019) [23]	Price index analysis	IaaS	Functional	3	52	4
Wu et al. (2018) [24]		IaaS	Both	48	199	5
Mitropoulou et al. (2016) [5]		IaaS	Functional	6	2742	26
Mitropoulou et al. (2017) [6]		IaaS	Both	17	806	23
Liagkou et al. (2022) [7]		CaaS	Both	18	640	6
Patel et al. (2022) [25]	Simple bundle comparison	IaaS	Both	10	4	3
Bari et al. (2015) [26]		IaaS	Functional	9	27	3
Rajput et al. (2023) [27]		IaaS	Both	23	4	3
Current work	Comparison based on price index analysis	IaaS/ CaaS/ PaaS	Both	11/ 15/ 13	589/ 640/ 806	6

customers have the option to adjust each feature according to their specific needs, thereby customizing their bundle and potentially changing the overall price each time.

Furthermore, it is critical to keep in mind that all providers charge network usage fees in addition to the price of the customer's cloud bundle. These fees are not included in Fig. 1 because the primary objective of our study is to explore the impact of features that are bundled within the price rather than those independently charged. Network usage fees are based on the amount of data transmitted and are generally similar across all services and providers. Consequently, they do not vary significantly and do not influence the pricing structure in a way that differentiates providers. Thus, network usage is not a primary consideration in our comparison of cloud service features, which focuses on hidden costs within the bundled price rather than separately billed items.

3.1 IaaS

Infrastructure-as-a-Service (IaaS) offers infrastructure such as virtual machines, storage, and an operating system. However, users manage and provision hardware and install applications. In exploring in depth the IaaS cloud services' resources, functional and non-functional features were chosen, offering a better description of the resource's functionality [6]. These features are presented below.

- *CPU*, defines the number of vCPUs.
- *RAM*, defines the size of memory in Gigabytes (GBs).
- *Storage*, is described by the capacity of a cloud resource.
- *Disk-Type*, is the type of storage disk. It refers to standard or solid-state disk (SSD) type.
- *Instance Type*, refers to the type of instance the customers create so as to run their applications. It indicates if the instance is spot, shared, dedicated, or isolated.
- *Payment Option*, describe the potential payment-based models: nothing upfront, partial amount upfront, or all amount upfront payment.
- *Term length*, defines if the cloud user chooses a usage commitment that leads to discounts.
- *Region*, indicates the geographical location where the resources are hosted.
- *Regional Redundancy*, defines if a resource is regional or zonal. Regional resources can be used by any resource in that region, while zonal resources can only be used by other resources in the same

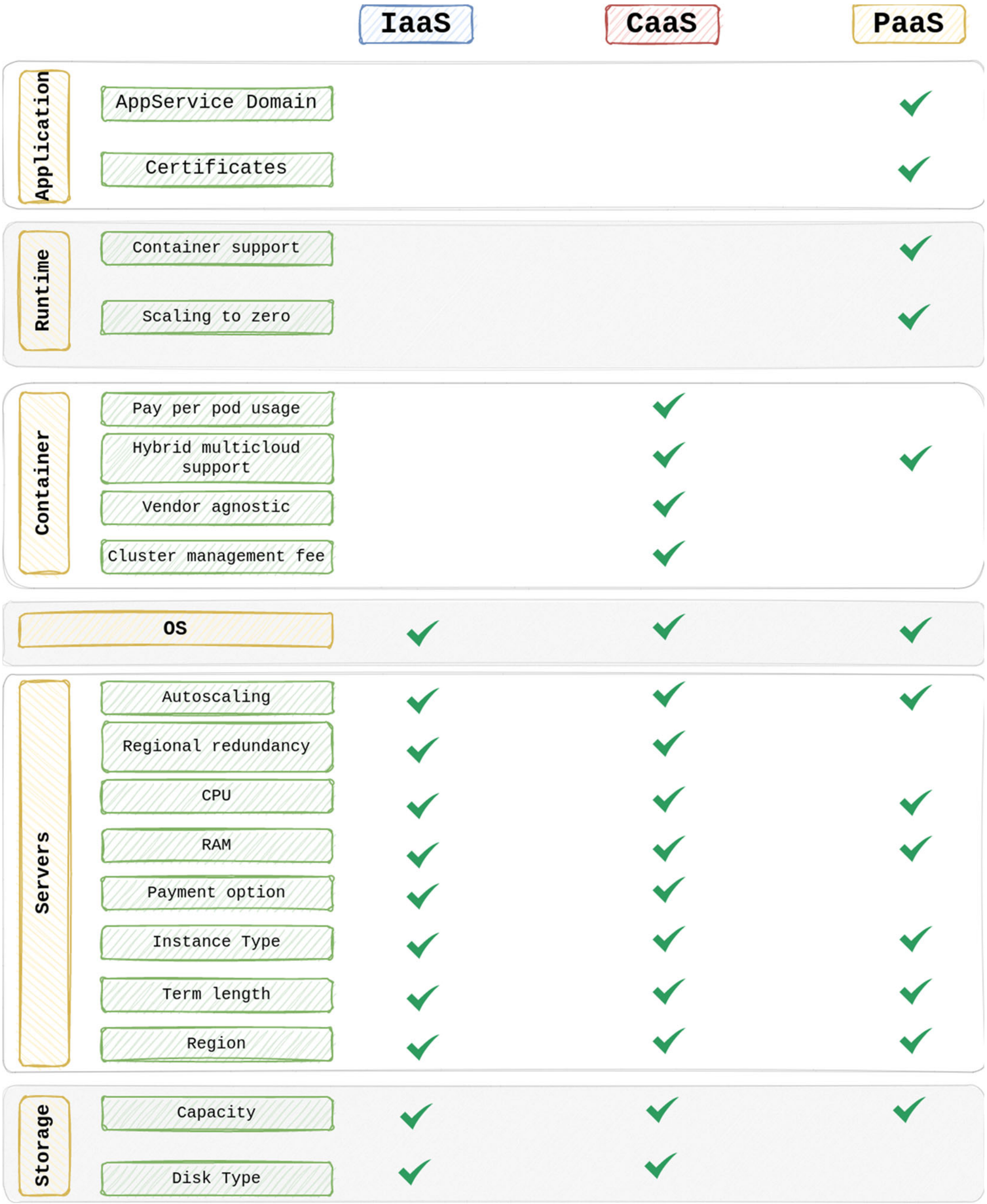


Fig. 1 Bundle features for IaaS, CaaS and PaaS service types

zone.

- *Auto-scaling*, ensures that resources are sufficient. Horizontal scaling refers to adding instances, whereas vertical auto-scaling refers to adding more or faster CPUs, memory, etc. to an existing instance.

Moreover, cloud virtualization environments need an operating system to manage the operation of virtual infrastructure. Cloud users can choose between an unlicensed and a commercial Linux or Windows-licensed OS.

3.2 CaaS

All the above features describe IaaS services, but the majority of them are also used by CaaS and PaaS services. Container-as-a-Service (CaaS) is one step up on the spectrum of cloud services. They offer container engines and orchestration functionality, mainly using Kubernetes, running on container hosts that are grouped into clusters [28]. CaaS providers bundle the container engine, orchestration tools, and underlying compute resources into a unified service. The well-known CaaS solutions that are studied in the current paper are Amazon Elastic Container Service (ECS) [29], Google Kubernetes Engine (GKE) [30], and Microsoft Azure Kubernetes Service (AKS)[31]. The following features are used to describe containerization in the cloud.

- *Pay per container usage* declares if the user is charged based on the actual container usage capacity instead of the underlying virtual machine capacity.
- *Hybrid and multi-cloud support* enables users to provision and manage containerized clusters of nodes running on multiple cloud providers or on-premise.
- *Cluster management fee* is a flat fee for cluster management, irrespective of cluster size and topology.
- *Vendor agnostic* declares if the container orchestration framework is based on the Kubernetes orchestration platform or another container orchestration platform.

3.3 PaaS

Platform-as-a-Service (PaaS) comes after IaaS and CaaS on the cloud services spectrum. PaaS is a cloud-based hosting model that focuses on application-level code deployment. It accesses IaaS resources but also provides a runtime environment for developers that they can build upon and use to create customized applications. Examples of PaaS solutions include AWS Elastic Beanstalk [29], Microsoft Azure Web Apps [31] and Google App Engine [30]. In the current paper, the three aforementioned PaaS offerings are studied, which are described by the following features:

- *Container support* indicates if the runtime environment can run container-based applications.
- *Scaling to zero* is a critical requirement for PaaS services. It indicates that when there are no inbound requests for an application, no instances of the application are running.
- *Certificates* determine the type of certificate to create, either a free certificate without additional charge from the provider, a standard certificate, or a wildcard certificate.
- *AppService Domain* indicates if the user is willing to buy a service domain and assign DNS names to the application.

4 Hedonic Price Indices for Cloud Services Pricing

The hedonic pricing method, also referred to as hedonic regression, makes the assumption that a product's or service's price depends on each of its unique characteristics. This method helps to measure how specific characteristics impact the total cost of a product or service [16]. We use the hedonic pricing model in the cloud to explore how functional and non-functional features, as indicated in Fig. 1, affect the total price of a particular service bundle. We construct hedonic price indices for each type of cloud model (IaaS, CaaS, and PaaS) and for each of the top three providers to emphasize the contribution of each entity in shaping the final price.

4.1 Hedonic Function

A hedonic function $f(X)$ is a relation between a number of the products' features and the corresponding prices, as presented in (1).

$$P_i = f(\mathbf{X_i}) + e(\mathbf{X_i}) \tag{1}$$

where P_i is the price of a product or service i and X_i is a vector of features related to the specific product or service, and e is the residual error. The simplest hedonic function form is linear regression using OLS (Ordinary Least Square), as presented in (2).

$$P = b_0 + \sum_{i=1}^{n} b_i X_i + e(\mathbf{X}) \tag{2}$$

where b_i are the estimated regression coefficients, describing the relationship between each independent variable X_i and the dependent variable P, indicating the prices charged and paid for an increment of one unit of the corresponding characteristic. However, the adoption of this form could create substantial errors, and some cloud features that are denoted as categorical variables cannot be denoted. Therefore, a semi-log hedonic pricing model, as shown in (3), was chosen as the most appropriate. Semi-log hedonic models are often preferred for their ease of estimation. Standard errors and statistical tests are easily computed [32].

$$\log P = b_0 + \sum_{i=1}^{n} b_i x_i + \sum_{t=1}^{k} \delta_i d_i + e(\mathbf{X}) \tag{3}$$

where $\log P$ is the log price of cloud service, x_i is the log-transformed values of the continuous variables, b_i are the estimated regression coefficients of the continuous variables, d_i are the categorical variables, and δ_i denote the estimated regression coefficients of the categorical variables. It is important to mention that b_0 represents the constant term of the equation, which accounts for the baseline level of the dependent variable that is not explained by the independent variables in the model. Equation 3 constitutes the basis of the pricing model analyzed in this paper.

In the context of a hedonic function, the constant term b_0 represents the implicit cost of a hypothetical good for which all explanatory variables have zero values. This indicates that b_0 captures the impact on price of all factors not accounted for by the hedonic model. These variables could be unobserved product features or market conditions not captured by the explanatory variables. In the hedonic literature, like in [33], referred to as the *baseline price*. The constant term b_0 can be used to compare the prices of various products on the market, even if their explanatory variables have distinct values.

To make a stable price index, the P-values and variance inflation factors (VIF) of each input variable are calculated. This makes sure that each variable has a clear and distinct relationship with the dependent variable and that the regression model for each variable is strong and reliable. Although the VIF is a valuable measure of multicollinearity, they primarily relied on the P-values to guide their variable inclusion decisions. In particular, they employed a significance threshold of 0.05 to determine whether each input variable made a significant contribution to the model, and they eliminated variables that did not meet this criteria. A stable price index that precisely reflected the relationship between input features and the dependent variable was produced, employing both the VIF and P-values in their analysis.

The variance inflation factor (VIF) is a method for quantifying multicollinearity, which reduces the precision of the estimated coefficients and weakens the rigor of the regression model [34]. It is determined using:

$$\mathrm{VIF}_i = \frac{1}{1 - R_i^2} \tag{4}$$

where R_i^2 is the coefficient of determination obtained when X_i is regressed on all other input variables in the model. As a general rule of thumb, a VIF value of less than 3 is considered to indicate no or low multicollinearity, while a VIF value of 4 to 9 indicates moderate multicollinearity. A VIF value of 10 or greater indicates strong multicollinearity, and it is generally recommended to remove the corresponding variable from the model. However, it is important to note that this rule of thumb is not a hard-and-fast rule, and the specific threshold for VIF that indicates multicollinearity can vary depending on the specific context and goals of the analysis.

4.1.1 Incorporation of Categorical Variables

One-Hot Encoding is a common method for incorporating categorical variables into statistical models. This requires the creation of dummy variables for each type within the categorical variable [35]. Each observation would have a 1 in the column corresponding to its selected option and 0s in the remaining columns. One-Hot Encoding is a popular technique because it permits the inclusion of categorical variables in regression models, which typically require numeric input variables. This method also avoids imposing a linear relationship between categories, a common presumption in other encoding techniques such as ordinal encoding [36]. In this research, we used it to incorporate categorical variables into this study's regression model. It generated dummy variables for each categorical variable and incorporated them into the model as input variables. At the same time, it enabled us to analyze the impact of categorical variables on the outcome variable while controlling for the possible effects of other variables.

A base case is a reference point or baseline that is used for comparison or analysis. In the context of categorical variables with more than two levels, the base case is the standard against which other levels are compared. These variables are essential for accurately modeling the relationship between the features of cloud services and their prices. By incorporating both functional and non-functional features into the model, a deeper comprehension of the factors that affect the pricing of cloud services can be gained. In addition, by carefully considering the base cases for categorical variables with more than two levels, researchers can ensure that their model is accurate, reliable, and reflects the true nature of the relationships between cloud attribute values and their prices.

4.2 Data collection and data features types

There are numerous cloud service providers offering the aforementioned cloud service types–IaaS[1], CaaS[2], and PaaS[3]– on the public cloud market. The authors obtained data manually using the official resources calculator on each service's website in September 2022. The providers chosen are Amazon, Microsoft, and Google, as they are the top three cloud providers and together obtain almost 65% of the cloud market [37]. These combinations of different features are numerous and consist of each bundle as depicted in Fig. 2. The IaaS dataset consists of 589 bundles, whereas the CaaS dataset includes 640 bundles, and finally, the PaaS dataset contains 806 bundles.

Specifically Fig. 2 illustrates the structural breakdown of a Cloud Service type. The service type is delineated into multiple bundles, each comprising various features. For instance, Bundle 1 includes several features that cumulatively determine the bundle's price per hour. Similarly, other bundles contain different combinations of features, each leading to a unique price per hour. This hierarchical structure showcases the component-based pricing model inherent in cloud service offerings, allowing for a detailed analysis of pricing strategies based on individual features within each bundle.

These data sets contain information on the various cloud service bundles offered by each service provider, including pricing and features. The features that describe cloud services are functional and non-functional. The functional features have numerical values, whereas the non-functional features are denoted by categorical values. Table 2 summarizes the value set for each attribute used to create price indices.

4.3 Experimental Flow

The experimental procedure consisted of a sequence of stages, as depicted in Fig. 3. Details on implementation and assumptions for calculating a hedonic index may be found in [7]. The entire set of features was then subjected to a regression procedure, and the resulting coefficients were exported. In addition to expert domain knowledge, *p-values* and *variance inflation factor*(VIF) were considered to select statistically significant values. The coefficients were then used to calculate price indices for each type of technology (IaaS, PaaS, and CaaS). In addition, price indices were calculated for each provider, based on a subset of the data, for each technology type individually. This procedure ensured that the most influential factors were identified and that the aggregate impact of each technology type and provider on the final price index was calculated.

5 Results

This section describes how to construct a price index for each cloud service type—IaaS, CaaS, and PaaS—and for each individual provider. Importantly, these price indices are formulated on a cost-per-hour basis, providing a standardized framework for comparison across services and providers. By analyzing the resulting price indices, we uncover insights into the pricing strategies employed by providers without the confounding factor of variable usage times. Cloud service consumers and decision-makers can leverage these findings to make informed decisions and optimize their cloud service expenditures, confident in the comparability of the cost metrics. 5.3 Hedonic function for PaaS

[1] IaaS dataset https://raw.githubusercontent.com/gfragi/cloudCasestudy/master/datasets/iaas_data.csv.

[2] CaaS dataset https://raw.githubusercontent.com/gfragi/cloudCasestudy/master/datasets/caas_data.csv.

[3] PaaS dataset https://raw.githubusercontent.com/gfragi/cloudCasestudy/master/datasets/paas_data.csv.

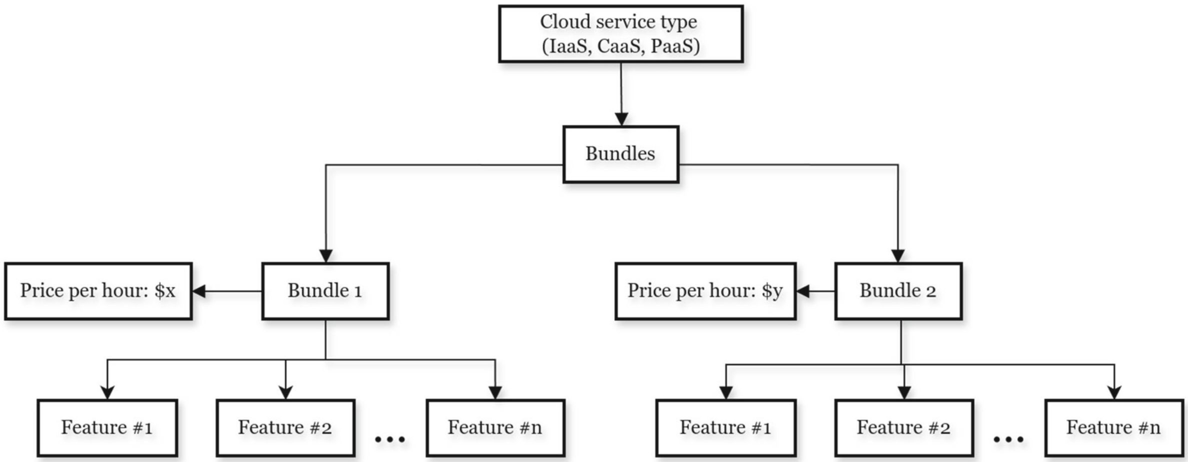


Fig. 2 Decomposition of cloud service types in terms of bundles and their features

5.1 IaaS Hedonic Function

In the hedonic pricing model, cloud features serve as independent variables, with functional features represented by numeric variables and non-functional features by categorical variables. The base cases for categorical variables with more than two levels for IaaS are displayed in Table 3.

Table 2 Functional and non-functional features

Features	Type	No. Values	Values
CPU (Cores)	N	7	2, 4, 8, 12, 16, 20, 32
RAM (GB)	N	8	4, 8, 16, 32, 48, 64, 128, 160
Storage (GB)	N	8	100, 200, 300, 400, 500, 800, 1TB, 2TB
Disk type	C	2	Standard, SSD
Operating system	C	3	Free OS, Linux licensed, Windows licensed
Instance type	C	3	On-Demand, Spot, Dedicated
Cluster management fee	C	2	Yes/No
Regional redundancy	C	2	Yes/No
Region	C	6	Europe, US, South America, Asia, Australia, Africa
Auto-scaling	C	2	Horizontal, Vertical
Hybrid/multi-cloud support	C	2	Yes/No
Vendor agnostic	C	2	Yes/No
Pay per container usage	C	2	Yes/No
Term length commitment	C	3	0, 1, 3 [years]
Payment option upfront	C	3	No, Partially, All
AppService domain	C	2	No, Yes
Certificates	C	3	No, Standard, WildCard
Container support	C	2	Yes, No
Scaling to zero	C	2	Yes, No

The price index based on IaaS services is shown in (5).

$$\log P = -0.44 + 0.53 * (\log CPU) + 0.19 * (\log RAM) + 0.02 * (\log Storage) + 0.02 * (\log Disk\_type) - 0.10 * (\log TermLength) + 0.10 * (InstType_{Dedicated}) - 0.07 * (InstType_{Spot}) - 0.03 * (OS_{Free}) + 0.03 * (OS_{Win}) - 0.01 * (Region_{US}) \quad (5)$$

The analysis of the hedonic function results for IaaS services reveals that the functional features, such as CPU and RAM, have the greatest influence on the pricing policy, whereas storage has the least impact. Concerning the non-functional factors, the *Spot type* lowers the cost nearly as much as the *Dedicated* increases it, when compared to the *On demand* instances as the base case. In addition, the *Operating System* type also affects the price, with unlicensed (free) and Windows-licensed operating systems having corresponding and opposing effects. Moreover, the region is the non-functional parameter with the least influence on pricing, with the price being lower when an instance runs on *US Region* rather than Europe, which is the reference case. Overall, these results imply that the pricing policy for IaaS services is predominantly influenced by functional features, with other non-functional features having comparatively smaller effects. Regression analysis determines the statistically significant features based on the *p-value* ($p \leq 0.05$) and VIF value. Table 4 presents the statistically significant coefficients of independent variables within a regression model used to forecast the pricing of IaaS services from three specific providers: Google, Amazon, and Microsoft.

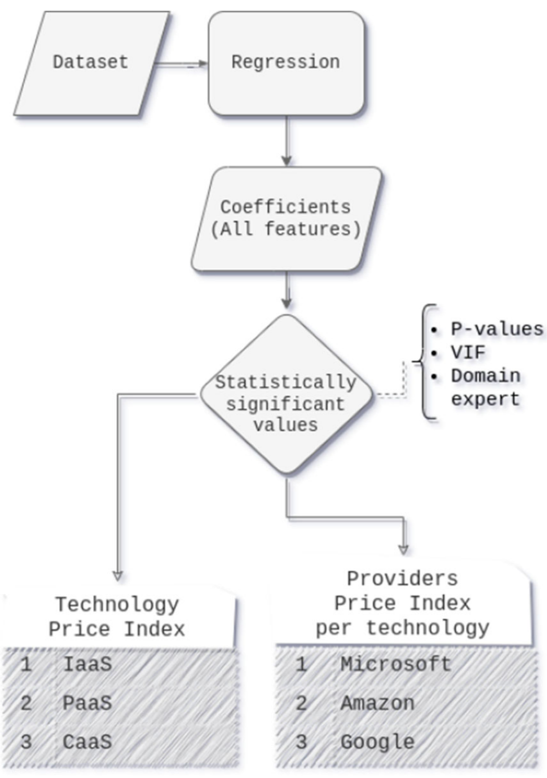


Fig. 3 Experimental flow

The model's log-transformed price variable is the dependent variable. When all other independent variables are equal to zero, the coefficient for the constant term indicates the expected log price. The coefficients for the other independent variables represent the expected change in log price associated with an increase of one unit in the corresponding independent variable, while all other variables are held constant.

5.2 Hedonic Function for CaaS

The price index based on CaaS services is shown in (6). The corresponding base cases for the categorical variables with more than 2 levels are presented in Table 5.

Table 3 Base case for IaaS categorical variables

Feature	Values	Appearances	
Payment method	No upfront	522	Base case
	Partial upfront	60	dummy
	All upfront	58	dummy
Instance Type	On demand	472	Base case
	Dedicated	109	dummy
	Spot	59	dummy
Operating System	Linux	310	Base case
	Free	185	dummy
	Windows	145	dummy
Region	Europe	244	Base case
	Asia	164	dummy
	US	146	dummy
	South America	29	dummy
	Australia	3	dummy
	Africa	3	dummy

$$
\begin{aligned}
\log P = &-0.55 + 0.61 * (\log CPU) + 0.11 * (\log RAM) + 0.05 * (\log Storage) \\
&-0.09 * (\log TermLength) + 0.08 * (ClusterFee) + 0.04 * (Autoscaling) \\
&+0.08 * (VendorAgnostic) + 0.05 * (MultiCloudSupp) + 0.08 * (PayPerCn) \\
&+0.05 * (OS_{Win}) + 0.09 * (InstTyp_{Dedicated}) - 0.04 * (InstTyp_{Spot}) \\
&-0.02 * (Region_{Asia}) - 0.02 * (Region_{US})
\end{aligned} \tag{6}
$$

According to the aforementioned price index data, CPU represents the most influential factor in the provider's pricing policy, comparable to that of IaaS. RAM comes next, although it has a significantly smaller impact on cost than the CPU. The non-functional CaaS features *Cluster fee*, *Vendor agnostic*, and *Pay per container* are all very close to RAM and each raise the cost identically. Moreover, the *Dedicated instance type* has nearly the same amount of weight as the previously stated objects, and its value in IaaS is essentially the same. In particular, when compared to *On demand* instances, it raises the cost by 0.9. On the other hand, if the customer purchases *spot* instances, the price is reduced by 0.4. Moreover, like in IaaS, the *TermLength* is the most significant cost-cutting element, while the *US Region* is the least. Furthermore, *Storage*, *Multicloud Support*, *Autoscaling*, and *Windows OS* all have the same effect on pricing, with the latter indicating that Windows containers are more expensive than Linux containers. Thereby, it is revealed that the pricing impact of common features between IaaS and CaaS is nearly the same. This is pretty intriguing since it proves that cloud providers follow the same pricing pattern for both services in terms of the common elements. Besides that, it is observed that all of the CaaS add-on features are statistically significant and have a noticeable impact on cost.

In addition, Table 6 displays the coefficients that are statistically significant for the three CaaS providers. The dependent variable is the price logarithm. All suppliers have a negative coefficient for the constant component, indicating that there is a fixed cost associated with providing the service. In the subsequent section, researchers use this result to evaluate the baseline price for the various providers and technologies.

Table 4 Statistically significant IaaS coefficients

Features	All	Google	Amazon	Microsoft
const	−0.44	−0.73	−0.36	−0.55
CPU	0.53	0.52	0.37	0.24
RAM	0.19	0.11	0.31	0.60
Storage	0.02	0.17	−0.03	−0.06
Disk_type	0.02	0.03	0.03	0.02
Term_Length	−0.10	−0.03	−0.12	−0.14
OS_Windows	0.03	0.06	0.00	0.01
OS_free	−0.03	0.01	0.00	−0.07
Instance_Type_Dedicated	0.10	0.00	0.12	0.02
Instance_Type_spot	−0.07	−0.01	−0.07	0.00
Region_US	−0.01	−0.13	−0.02	0.02

5.3 Hedonic Function for PaaS

Continuous and categorical variables are used to represent the components that constitute an application-type PaaS service, which offers development and deployment environments for application services [38]. Table 7 indicates the base cases for categorical variables with more than 2 levels. The price index based on PaaS services is shown in (7).

$$\begin{aligned} \log P = &-0.16 + 0.31 * (\log CPU) - 0.01 * (\log RAM) + \\ &0.05 * (\log Storage) - 0.13 * (\log TermLength) + 0.05 * (OS_{Win}) \\ &-0.01 * (AppService) + 0.02 * (CnSupport) + 0.08 * (InstType_{Isolated}) \\ &+0.02 * (ScalingZero) - 0.01 * (Autoscaling) - 0.02 * (Cert\_stand) \\ &+0.05 * (Africa) + 0.04 * (Asia) + 0.06 * (Australia) + 0.03 * (US) \\ &+0.06 * (S.America) \end{aligned} \tag{7}$$

Table 8 displays the statistically significant values for the three cloud service companies' PaaS models: Google, Amazon, and Microsoft.

The PaaS pricing index results demonstrate once again that CPU is a significant price element for cloud costs. In addition, the *Term Length* commitment represents the biggest source of cost savings, as is the case with IaaS and CaaS. Another comparison of PaaS findings with IaaS and CaaS results shows that the *Isolated instance* type has a similar weight to the other two services and also contributes to the highest cost rise after CPU.

Together with the aforementioned, the *Storage* and *Windows OS* have the exact same impact with respect to the other two services, with the latter proving that Windows is more costly than commercial Linux, even for PaaS. Furthermore, the influence of *Region* attribute on PaaS pricing is worth highlighting, demonstrating that Europe is the cheapest region for hosting PaaS services while Australia is the most expensive. Moreover, *Container support* and *Scaling to zero* are two features that have the least influence on price increases, reflecting the providers' approach in response to the rise of container-based solutions. Lastly, *AppService domain* and *Autoscaling* are also included, albeit with the lowest effect, implying that they are standard parts of PaaS offerings.

6 Comparison Results - Managerial Implications

6.1 Providers Comparison per Cloud Service Type

In this section, we explore the influence of cloud features on the shaping of the cloud price. Therefore, we compare the statistically important functional and non-

Table 5 Base case for CaaS categorical variables

Feature	Values	Appearances	
Payment Method	No upfront	522	Base case
	Partial upfront	60	dummy
	All upfront	58	dummy
Instance Type	On demand	472	Base case
	Dedicated	109	dummy
	Spot	59	dummy
Operating System	Linux	310	Base case
	Free	185	dummy
	Windows	145	dummy
Region	Europe	279	Base case
	Asia	172	dummy
	US	155	dummy
	South America	34	dummy

functional features of each service rather than the total number of features as depicted in Fig. 1.

6.1.1 IaaS Providers Comparison

Comparing the price index results for the three leading providers, it is evident that all providers follow a pretty similar pricing pattern for IaaS services, as presented in Fig. 4. The associated values are provided in further detail in Table 4.

Below are the key points of the comparative analysis. We will first examine the similarities in pricing policies for IaaS across providers, highlighting the key aspects that are consistent.

- CPU and RAM are the most essential cost aspects of IaaS across all providers. This is reasonable consid-

Table 6 Statistically significant CaaS coefficients

Features	All	Google	Amazon	Microsoft
const	−0.55	−0.52	−0.22	−0.55
CPU	0.61	0.68	0.53	−0.08
RAM	0.11	0.02	0.19	0.90
Storage	0.05	0.07	0.03	−0.04
Multicloud_Support	0.05	0.01	0.04	−0.03
Pay_per_Container	0.08	0.10	0.00	0.00
Vendor Agnostic	0.08	0.00	−0.01	0.00
Cluster_mgmt_fee	0.08	0.11	−0.21	0.10
Instance_Type_Dedicated	0.09	0.00	0.13	0.06
Instance_Type_Spot	−0.04	−0.02	−0.07	0.00
Autoscaling	0.04	0.00	0.00	−0.23
OS_Windows	0.05	0.04	0.04	0.03
Term_Length	−0.09	−0.05	−0.09	−0.14
Region_Asia	−0.02	0.05	−0.07	0.02
Region_US	−0.02	−0.03	0.08	0.02

Table 7 Base case for PaaS categorical variables

Feature	Values	Appearances	
Instance Type	Shared	466	Base case
	Dedicated	244	dummy
	Isolated	96	dummy
Certificate	No	710	Base case
	Wildcard	50	dummy
	Standard	46	dummy
Region	Europe	162	Base case
	Asia	142	dummy
	US	270	dummy
	South America	117	dummy
	Australia	79	dummy
	Africa	36	dummy

ering that the CPU and memory are two of the most energy-consuming components of cloud computing infrastructure [39]. Consequently, energy prices dictate their costs, rendering them the most expensive components of a cloud service.

- The commitment term has a significant impact on pricing across all three providers, with Amazon and Microsoft offering more substantial discounts for longer-term commitments compared to Google.
- The impact of disk type is also consistent across Amazon and Google policies, with slight variations observed in Microsoft's approach.

Subsequently, we will analyze the variations in pricing strategies among providers, emphasizing the distinct factors that differentiate them.

- The influence on price for CPU and RAM differs greatly amongst providers such as Google and Microsoft, which are diametrically opposed.
- A relatively similar concept applies for the storage attribute, although at a lower level. This impact mismatch is driven by variations in data center equipment and locations between providers.
- Instance type has a determinant influence on pricing after CPU and RAM. The dedicated and spot instance types have a considerably greater influence on increasing or reducing IaaS prices at Amazon than the other two.
- The Windows operating system has a greater impact on pricing at Google, a lesser at Microsoft, and no effect on Amazon's pricing strategy.
- In Google, the price decreases significantly when an instance is operated in the US region compared to Amazon's instance running in the same region. Conversely, there is a price increase when a Microsoft instance is utilized in the US region.

6.1.2 CaaS Providers Comparison

The CaaS comparison between providers is shown in Fig. 5, whereas the associated values are provided in further detail in Table 6.

Similar to IaaS, it is clearly demonstrated that all providers follow the same pattern for CaaS pricing. *CPU* and RAM are the two major elements influencing CaaS pricing, precisely similar to what happens in IaaS. This is reasonable given that CaaS is built on IaaS's underlying compute resources. As a consequence, the IaaS results influence the CaaS outcome.

However, there are several similarities and differences. In terms of similarities, the analysis reveals the following:

- In Amazon and Google, the CPU significantly increases the price, while in Microsoft, it leads to a decrease in price.
- The storage attribute follows a similar concept. Amazon and Google storage increase prices, whereas Microsoft leads in decreases. In both IaaS and CaaS, Google offers storage with an additional charge, whereas Microsoft includes it as part of the company's standard service provision.
- *Vendor agnostic* and *multi-cloud support* features have a nearly identical effect on CaaS pricing for all providers. These features give enterprises the agility and flexibility to deploy containers anywhere in the cloud or on-premise infrastructure. Organisations may therefore future-proof their container deployment strategy by avoiding vendor lock-in and ensuring business continuity in the event of a cloud provider disaster [40].
- The impact of *cluster management fee* on price is roughly the same among providers besides Amazon. This expense is connected with provider services such as managing and maintaining the

Table 8 Statistically significant PaaS coefficients

Features	All	Google	Amazon	Microsoft
const	−0.16	−0.08	−0.02	−0.07
CPU	0.31	0.21	0.21	0.20
RAM	−0.01	0.05	0.01	0.18
Storage	0.05	0.06	0.02	−0.01
Term_Length	−0.13	0.00	−0.06	−0.16
OS_Windows	0.05	−0.02	0.00	0.05
AppService_Domain	−0.01	−0.02	−0.01	−0.01
Container_support	0.02	−0.01	0.01	−0.07
Instance_Type_Isolated	0.08	0.00	0.01	0.14
Scaling_to_zero	0.02	0.01	−0.03	0.00
Autoscaling	−0.01	0.01	−0.02	0.07
Certificates_Standard	−0.02	0.00	0.00	−0.05
Region_Africa	0.05	0.00	0.01	0.07
Region_Asia	0.04	0.02	0.04	0.06
Region_Australia	0.06	0.05	0.00	0.04
Region_South_America	0.06	0.06	0.10	0.03
Region_US	0.03	0.04	0.05	−0.02

underlying infrastructure required to host containers. Offering cluster administration at a discount appears to be Amazon's strategy, which might bring down the cost.

- The *Pay-per-Container* remains relatively consistent across providers. However, there is a slight differentiation observed at Google. It seems that Google seeks to provide a serverless experience to its customers by offering this option, so that they may only pay for the containers they really use.
- *Autoscaling* appears to be a standard service offering by the company. In Google and Amazon, it does not affect the pricing policy of CaaS, whereas in Microsoft, it leads to a notable decrease in pricing. Moreover, in the IaaS model, Autoscaling had no impact on pricing as it was not statistically significant.
- Similar to the IaaS model, the *Windows operating system* leads to an increase in the final price.
- The *term length* demonstrates a negative coefficient across the three providers, indicating that the price of the service decreases as the duration of the contract increases. A similar observation was made in the IaaS model as well.
- *Dedicated* instances result in price increases, whereas *spot* instances lead to price decreases. The same trend was observed in IaaS, and this is justifiable as CaaS is constructed upon the underlying compute resources of IaaS.

However, there are distinctions among the providers in their pricing policies for CaaS.

- At Microsoft, RAM and CPU have opposing effects. CPU decreases in price, whereas RAM plays the most determinant role in pricing, leading to an increase in the final price. In contrast, at Google and Amazon, CPU and RAM have similar impacts on pricing.
- In Microsoft, prices rise when instances run in the US and Asia regions and are more expensive than in Europe. In Google, prices increase when instances run in the Asia region but decrease when they run in the US. Conversely, in Amazon, an instance in the Asia Region can be cheaper than in the US Region. Comparing to the Iaas model, Google is cheaper in the US.

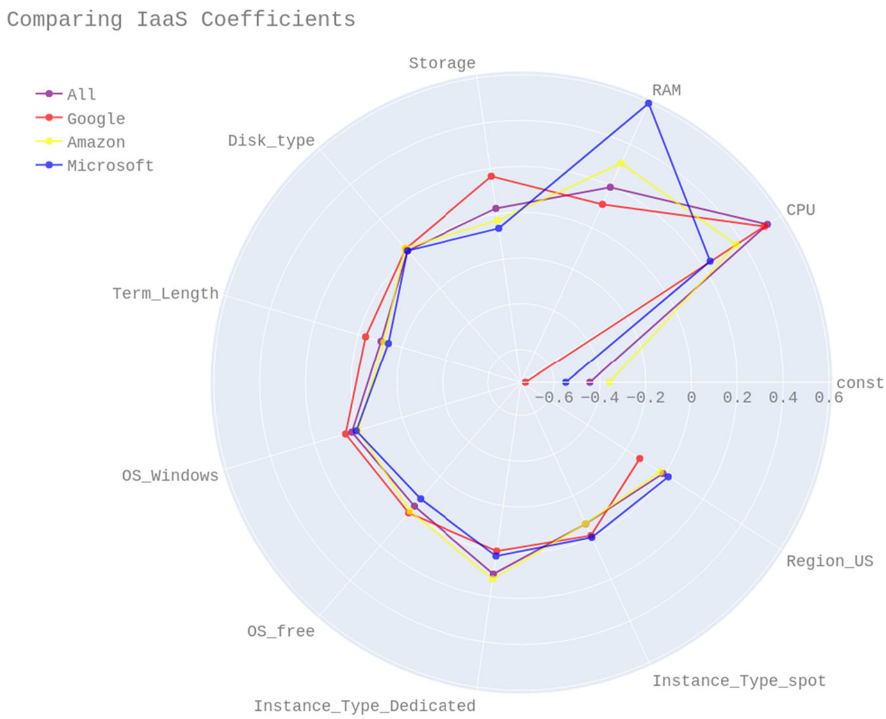


Fig. 4 IaaS providers comparison

6.1.3 PaaS Providers Comparison

Figure 6 presents the comparative results between the three leading cloud providers for PaaS.

We notice that there is a differentiation between PaaS, and IaaS and CaaS. Except for the CPU, nearly every other attribute has a different pricing impact depending on the provider, with Microsoft exhibiting the most significant discrepancies compared to the others. Specifically, the common pricing trends for PaaS across providers are the following:

- *CPU* is the primary factor determining PaaS costs for all three providers, and the effect of *CPU* on pricing seems to be a perfect match across providers. In IaaS and CaaS services, customers are charged depending on how long an instance remains allocated, regardless of whether the instance is fully utilized or not. However, PaaS services are charged based on how many CPU cycles a customer's application consumes. Thus, it is evident that the more processing power an application requires, the more expensive it will be to run on a PaaS platform.
- Across the providers, RAM increases in price; however, it has far less impact compared to IaaS and CaaS.
- *AppService Domain* contributes to price decreases across providers.
- *Term Length* factor results in price reduction across all providers and, particularly for Microsoft, may influence cost approximately as much as CPU. Similar findings were noted in the IaaS and CaaS models, highlighting the significant role of this feature in pricing strategies overall.
- The *operating system* consistently results in price increases across all providers. It is entirely reasonable that this particular observation was validated across the three models. However, it is one of the factors having different pricing impacts between Microsoft and the other two. In PaaS, the operating system is managed by the cloud provider. The operating system chosen determines the infrastructure required to serve PaaS applications, which might affect the overall cost. PaaS providers offer a variety of supported programming languages, frameworks, and run-time environments, which can limit the choice of platforms available and exposure pricing. The Azure platform offers a PaaS service for applications with a specific focus on the .NET framework [41], which may influence the differential impact on the final price.

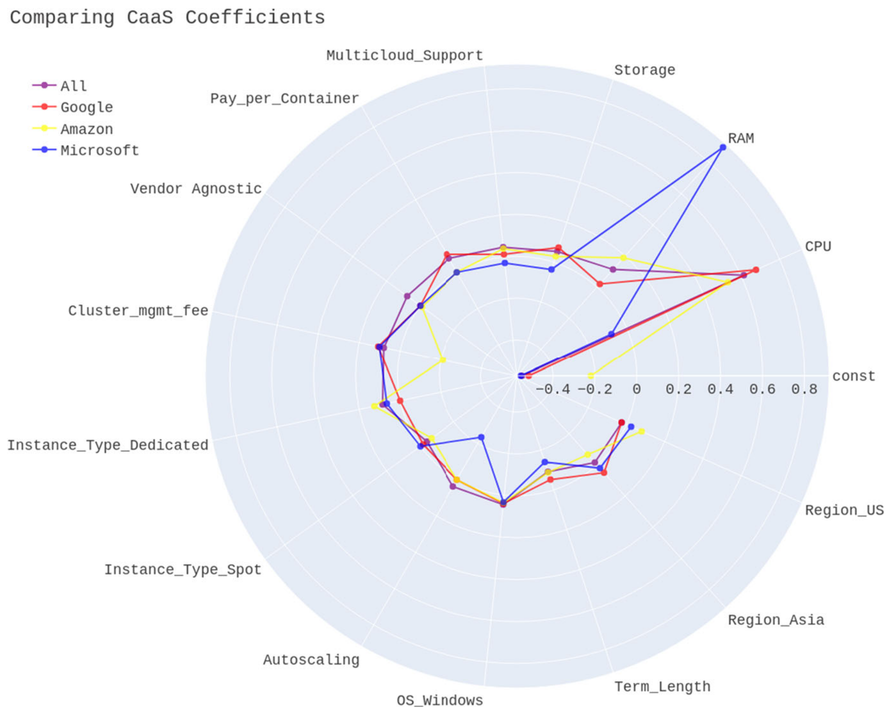


Fig. 5 CaaS providers comparison

- Instance type *Isolated* contributes to price increases across providers, particularly Microsoft.
- *Certificate* is not offered with an additional charge across all providers, highlighting the importance of application security.
- *AppService domain* lowers the PaaS price across the providers.

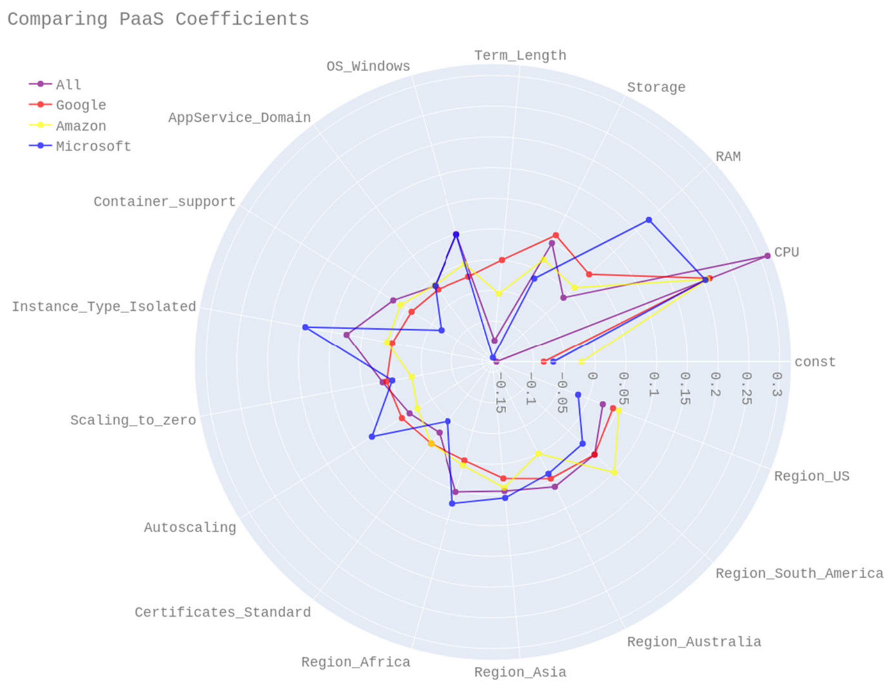
Pointing to the variations in PaaS pricing policies across providers, we emphasize the following:

- Google and Amazon Storage increase the price, whereas Microsoft stands out with a decrease in pricing. Microsoft, in three models, considers storage as part of the company's standard service provision.
- *Auto-scaling* increases the price in Google and Microsoft while decreasing in Amazon. In CaaS, the three providers offer Autoscaling as the company's standard service provision. However, in PaaS, Google and Amazon offer it with an additional charge, while Microsoft maintains the same policy.
- *Scaling to zero* decreases the price at Amazon and leads to an increase at Google. Additionally, it has no impact on Amazon's policy.
- *Container support* is a feature that mainly lowers PaaS prices in Google and particularly Microsoft, whereas increasing Amazon prices.
- In terms of region, Platform as a Service (PaaS) diverges from the other models, as all factors related to region exhibit statistical significance, specifically in PaaS. In Amazon and Google, the South America region offers the greatest increase in price, while the Microsoft Africa region leads in a greater price increase.

6.2 Cloud Service Comparison per Service Type

All three cloud service types - IaaS, CaaS, and PaaS - are described using a core of functional and non-functional features. In order to explore whether such common features have the same importance across different categories, a comparison of the three service models using common and statistically important features is performed. The statistically significant features of each cloud service type are presented in the appropriate Tables of Sections 5.1, 5.2, and 5.3, where we explain how the p-value assists in keeping the features that genuinely effect the pricing per hour for each hedonic function. To represent the intersection of statisti-

Fig. 6 PaaS providers comparison



cally important features among IaaS, PaaS, and CaaS price indices, as extracted by the regression analysis, a Venn diagram was developed as presented in Fig. 7.

Figure 8 demonstrates the comparison between service categories for the whole dataset and by provider. Unlike Microsoft, Amazon and Google follow near-identical price patterns for the three cloud services, which fits the overall pricing approach given in Fig. 8a.

Below, we present the influence of the core features of IaaS services on the pricing of both PaaS and CaaS. Especially, we emphasize the shared pricing patterns as follows:

- CPU emerges as the predominant cost factor across all providers and specifically in Amazon and Google's services, while its significance is relatively lower in Microsoft's offerings. Particularly in Amazon and Google's CaaS platforms, CPU plays the most crucial role in pricing, followed by IaaS and PaaS. In Microsoft, CPU has a comparable impact on IaaS and PaaS prices and a lower impact on CaaS.
- RAM plays a critical role across all providers, with its influence most pronounced in Microsoft's pricing strategy. Specifically, RAM significantly impacts CaaS pricing, followed by IaaS and PaaS. At Amazon, RAM exerts the greatest influence on IaaS pricing, while its impact on PaaS pricing is minimal. Meanwhile, at Google, RAM strongly affects IaaS pricing, with similar impacts on PaaS and CaaS pricing.
- Storage has a comparable influence on the cost of cloud models across all providers. However, Google storage has a greater effect on the price of IaaS compared to PaaS and CaaS.
- Term length and the operating system Windows similarly affect pricing across all cloud models and providers.

Nevertheless, the following distinctions are evident:

- For most providers, the US region doesn't affect pricing. However, both Amazon and Google offer IaaS at a lower cost in the US region compared to Europe, with a higher rate discrepancy than other models.
- As previously stated, the constant term signifies the influence of cloud features that were not explicitly considered in constructing the price indices on pricing policy. In Google and Amazon, the IaaS service has the lowest constant among services, while PaaS seems to have the highest. In Microsoft, IaaS and CaaS have similar constant terms, whereas in PaaS, the constant has a higher value. These findings are reasonable given that IaaS delivers the most fun-

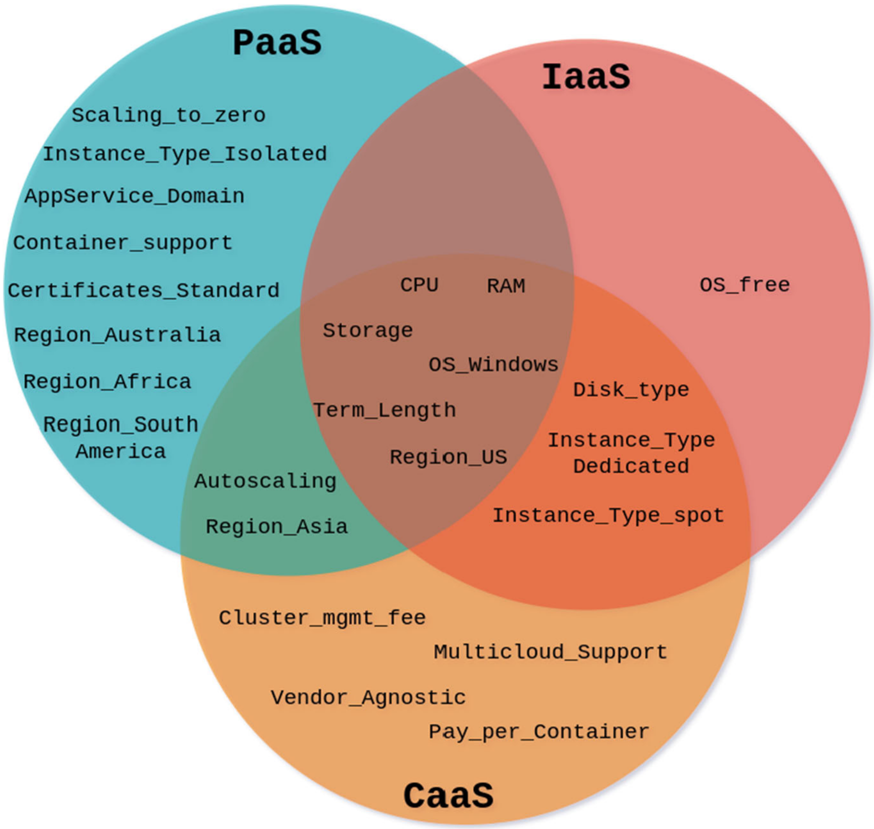


Fig. 7 Statistically significant features per service type

damental and basic level of managed cloud computing services. The higher-level services for CaaS and PaaS, which offer more advanced features and functionality, increase the commitment term.

7 Conclusions - Future Work

This work proposes a methodology to assess and compare pricing policies between IaaS, CaaS, and PaaS cloud service types and their main cloud providers. It is based on the construction and comparison of hedonic indices to: a) identify the features with a statistically important impact on the bundle price for each case; and b) provide a common function to facilitate comparison between different service categories and providers. The study's findings could help decision-makers optimize their cloud investments and make the most effective strategic decisions possible.

Based on our findings, all three cloud providers follow approximately the same pricing strategy for IaaS and CaaS, which reveals that CPU and RAM are by far the most vital cost factors, whereas the rest have a cumulative impact on the bundle price. Concerning PaaS, there is a distinction in pricing patterns between providers. Although CPU is still the key factor determining cost, almost every other attribute has a differential pricing impact depending on the provider, with Microsoft establishing a distinct pricing policy and having the most price variations compared to the other two. Although not a surprise, it is worth noting that as the complexity of the service bundle grows, incorporating additional features, the starting price for the service also increases.

The study utilized the hedonic price index methodology to analyze cloud service pricing, focusing on workload-specific requirements. This method involves giving varying weights to important features like CPU and RAM to analyze in detail how these aspects impact the cost framework of cloud services. Users can customize their assessment of cloud services to match the requirements of their workloads, enabling them to make well-informed decisions that cater to the specific needs of their computing jobs.

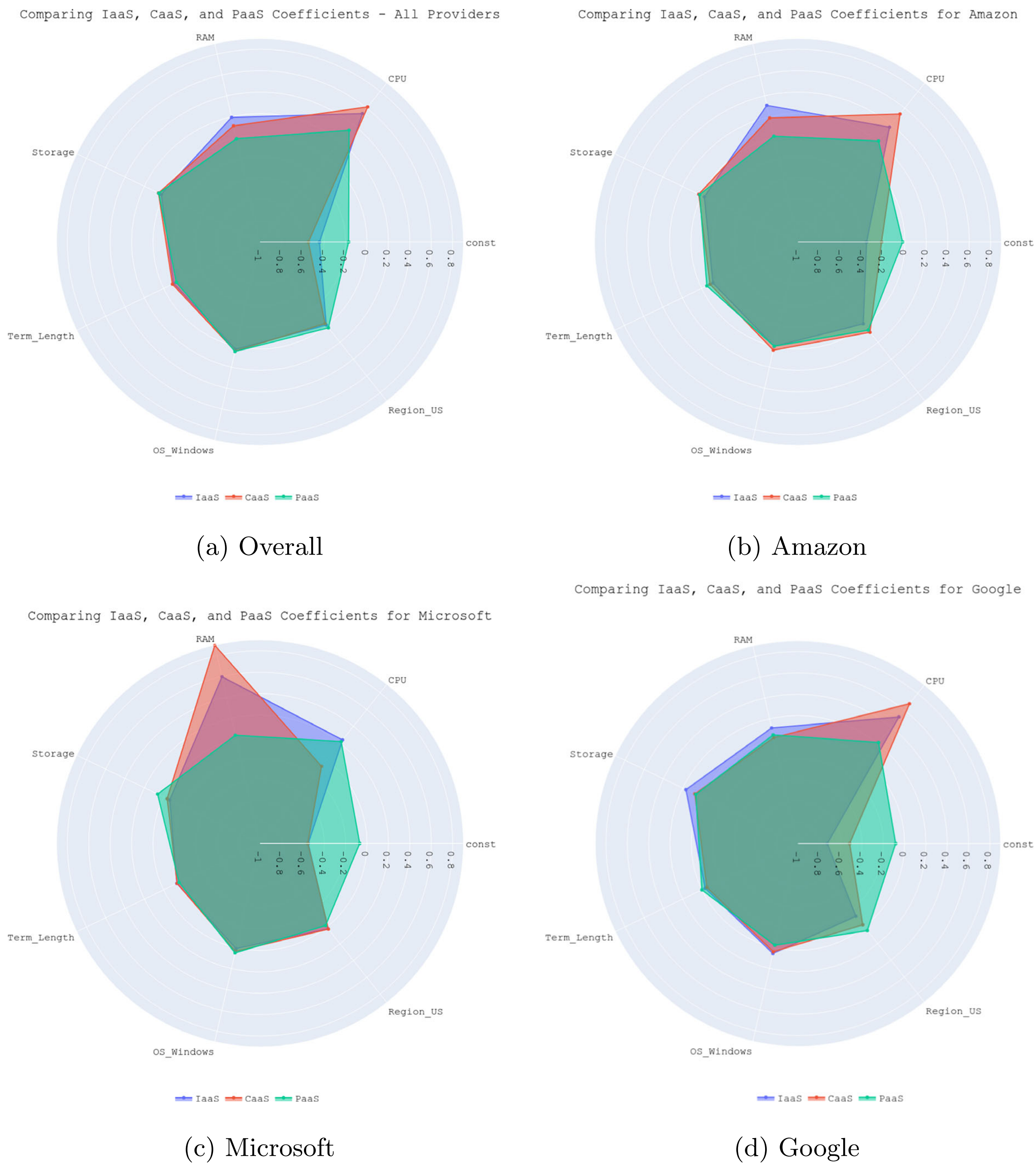


Fig. 8 Comparing cloud service types using the common features

The paper's careful approach intentionally deferred the creation of specific selection guidelines for cloud services to future research efforts. This choice recognizes the complex process of aligning specific pricing data with various client needs, requiring a combination of analytical precision and computing speed.

Looking ahead, it becomes evident that the integration of algorithmic decision-making tools, such as decision trees, holds considerable promise for enhancing the selection process of cloud services. Such methodologies could offer a structured and scalable framework for navigating the complex interplay between service pricing and user requirements, ultimately leading to the formulation of robust, data-driven guidelines for cloud service selection.

By providing a baseline for future comparisons of similar cloud technologies, this study is an important asset to the field of cloud computing because it sheds light on the pricing strategies implemented by major cloud providers and the factors that affect cloud costs. Further extensions and future research directions include the study of pricing for hybrid solutions such as sovereign clouds or pricing in serverless cloud models like Function-as-a-Service.

Author contributions V.L and E.F Contributed to the conceptualization and design of the study, data collection, and analysis. G.F and M.N. developed of the methodology, particularly in the application of the hedonic price index. Contributed to the review and editing of the manuscript. C.M. and A.T contributed to the conceptualization of the study and provided expertise on cloud computing economics, assisting in the interpretation of the data. All authors reviewed the manuscript.

Data Availability Data is provided within the manuscript as link in the footnote.

Declarations

Competing interest The authors declare no competing interests.

References

1. Soni, A., Hasan, M.: Pricing schemes in cloud computing: a review. Int. J. Adv. Comput. Res. **7**(29), 60 (2017)
2. Anodot.: State of cloud cost report (2022)
3. Cloudzero.: The state of cloud cost in (2024)
4. Foundry.: Insight into cloud computing trends
5. Mitropoulou, P., Filiopoulou, E., Michalakelis, C., Nikolaidou, M.: Pricing cloud iaas services based on a hedonic price index. Comput. **98**(11), 1075–1089 (2016)
6. Mitropoulou, P., Filiopoulou, E., Nikolaidou, M., Michalakelis, C.: Pricing iaas: a hedonic price index approach. In: International conference on the economics of grids, clouds, systems, and services, pp. 18–28. Springer (2017)
7. Liagkou, V., Fragiadakis, G., Filiopoulou, E., Michalakelis, C., Kamalakis, T., Nikolaidou, M.: A pricing model for container-as-a-service, based on hedonic indices. Simul. Model. Pract. Theory. **115**, 102441 (2022)
8. Mell, P., Grance, T., et al.: The nist definition of cloud computing (2011)
9. ITU-T.: 3500 information technology-cloud computing-overview and vocabulary (2014)
10. Hogan, M., Liu, F., Sokol, A., Tong, J.: Nist cloud computing standards roadmap. NIST Special Publication **35**, 6–11 (2011)
11. Kratzke, N.: A brief history of cloud application architectures. Appl. Sci. **8**(8) (2018)
12. Leymann, F., Breitenbücher, U., Wagner, S., Wettinger, J.: Native cloud applications: why monolithic virtualization is not their foundation. In: Cloud computing and services science (M. Helfert, D. Ferguson, V. Méndez Muñoz, and J. Cardoso, eds.), (Cham), pp. 16–40, Springer International Publishing, 2017
13. Kaushik, V., Bhardwaj, P., Lohani, K.: Game of definitions-do the nist definitions of cloud service models need an update? a remark. In: Futuristic trends in networks and computing technologies: select proceedings of fourth international conference on FTNCT 2021, pp. 653–666. Springer (2022)
14. Kaushik, V., Bhardwaj, P., Lohani, K.: Layered service model architecture for cloud computing. In: International conference on cloud computing, pp. 91–106. Springer (2022)
15. Laatikainen, G., Ojala, A., Mazhelis O.: Cloud services pricing models. In: Software business. from physical products to software services and solutions: 4th international conference, ICSOB 2013, Potsdam, Germany, June 11-14, 2013. Proceedings 4, pp. 117–129. Springer (2013)
16. Triplett, J.: Handbook on hedonic indexes and quality adjustments in price indexes: special application to information technology products (2004)
17. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., et al.: A view of cloud computing. Commun. of the ACM **53**(4), 50–58 (2010)
18. Moreira, A.S., Miers, C.C., Koslovski, G.P., Pillon, M.A., Gonzalez, N.M.: Energy-based cost model of containers provisioning on clouds. Int. J. Grid Util. Comput. **13**(6), 607–623 (2022)
19. Hinz, M., Koslovski, G.P., Miers, C.C., Pilla, L.L., Pillon, M.A.: A cost model for iaas clouds based on virtual machine energy consumption. J. Grid Comput. **16**, 493–512 (2018)
20. Aravantinos, E., Petre, K., Katsianis, D., Chipouras, A., Varoutas, D.: A hedonic approach to estimate the price evolution of ftth service: evidence from eu (2022)
21. Yun, S.H., Kim, Y., Kim, M.: Quality-adjusted international price comparisons of mobile telecommunications services. Telecomm. Policy **43**(4), 339–352 (2019)
22. Ahmad, W., Ahmed, T., Ahmad, B.: Pricing of mobile phone attributes at the retail level in a developing country: hedonic analysis. Telecomm. Policy **43**(4), 299–309 (2019)
23. El Kihal, S., chlereth, C., Skiera, B.: Price comparison for infrastructure-as-a-service.. In: ECIS, pp. 161 (2012)
24. Wu, C., Toosi, A.N., Buyya, R., Ramamohanarao, K.: Hedonic pricing of cloud computing services. IEEE Trans. Cloud Comput. **9**(1), 182–196 (2018)
25. Patel, A.R., Tiwari, R.V., Khureshi, R.A.: Comparative study of top cloud providers on basis of service availability and cost. IJFMR-Int. J. Multidiscip. Res. **4**(6) (2022)
26. Bari, I., Babu, S., Iqbal, M.M., Saleem, Y., Masood, Z.A.: Cost and performance based comparative study of top cloud service providers. Int. J. Comput. Sci. Inf. Secur. (IJCSIS) **13**(12), 172–177 (2015)
27. Rajput, A., Gupta, P., Ghodeshwar, P., Varma, S., Sharma, K., Singh, U.: Study of cloud providers (azure, amazon, and oracle) according to service availability and price. In: 2023

3rd International conference on pervasive computing and social networking (ICPCSN), pp. 1177–1188. IEEE (2023)
28. Bernstein, D.: Containers and cloud: from lxc to docker to kubernetes. IEEE Cloud Comput. **1**(3), 81–84 (2014)
29. Amazon.: Amazon elastic container service (2023)
30. Engine, G.K.: The most scalable and fully automated kubernetes service (2021)
31. Microsoft.: Pay as you go with azure (2022)
32. Elliott, G., Timmermann, A.: Handbook of economic forecasting. Elsevier (2013)
33. Greenstein, S., McDevitt, R.C.: The broadband bonus: estimating broadband internet's economic value. Telecomm. Policy **35**(7), 617–632 (2011)
34. Licht, M.H.: Multiple regression and correlation (1995)
35. Potdar, K., Pardawala, T.S., Pai, C.D.: A comparative study of categorical variable encoding techniques for neural network classifiers. Int. J. Comput. Appl. **175**(4), 7–9 (2017)
36. Huang, J., Ling, C.X.: Using auc and accuracy in evaluating learning algorithms. IEEE Trans. Knowl. Data Eng. **17**(3), 299–310 (2005)
37. Statista.: Google cloud, companies and products series (2022)
38. Gartner.: Platform as a service (paas) (2022)
39. Katal, A., Dahiya, S., Choudhury, T.: Energy efficiency in cloud computing data centers: a survey on software technologies. Cluster Comput., pp. 1–31, (2022)
40. Kansal, S., Singh, G., Kumar, H., Kaushal, S.: Pricing models in cloud computing. In: Proceedings of the 2014 international conference on information and communication technology for competitive strategies, pp. 1–5, 2014
41. Ferrer, A.J., Pérez, D.G., González, R.S.: Multicloud platform-as-a-service model, functionalities and approaches. Proced. Comput. Sci. **97**, 63–72 (2016)