

ARTIFICIAL STOCK MARKETS: A LITERATURE REVIEW

MERCADOS DE AÇÕES ARTIFICIAIS: UMA REVISÃO DA LITERATURA

Conceição Jacqueline X. B. de Castro ; <https://orcid.org/0000-0003-2461-1861>.
Universidade Presbiteriana Mackenzie

Eli Hadad Junior ; <https://orcid.org/0000-0003-2985-9241>
Universidade Presbiteriana Mackenzie

Michele Nascimento Jucá ; <https://orcid.org/0000-0002-8610-6193>
Universidade Presbiteriana Mackenzie

ARTIFICIAL STOCK MARKETS: A LITERATURE REVIEW

MERCADOS DE AÇÕES ARTIFICIAIS: UMA REVISÃO DA LITERATURA

ABSTRACT

This study aims to carry out a bibliometric analysis and systematic review of publications on artificial stock markets. The final sample consists of 52 articles. Such articles are obtained from the WoS and Scopus databases. Bibliometric analysis describes the frequency, citations and relationships between the main data of articles – authors, journals, keywords, etc. In addition, bibliometric laws are verified – Lotka (1926), Bradford (1934) and Zipf (1949). The bibliometric analysis is carried out using the Rstudio and Biblioshiny software. In turn, the systematic review seeks to identify knowledge gaps in the existing literature. This part of the research is still under development. The partial results of the research – resulting from the bibliometric analysis – indicate that the most prominent keywords are artificial stock market, genetic programming and agent-based modeling. Shu-Heng Chen appears as the most relevant author and co-author. When considering the institutions to which these authors are linked, it appears that the country with the largest number of published articles is China. However, the articles with the highest number of citations are linked to authors from institutions in the United States. The three bibliometric laws were confirmed.

Keywords: Artificial stock markets, Agent-based modeling, Genetic programming, Bibliometric analysis, Systematic review

RESUMO

Este estudo tem como objetivo realizar uma análise bibliométrica e revisão sistemática das publicações sobre os mercados de ações artificiais. A amostra final é composta por 52 artigos. Tais artigos são obtidos a partir das bases WoS e Scopus. A análise bibliométrica descreve a frequência, citações e relações entre os principais dados dos artigos – autores, periódicos, palavras-chave etc. Além disso, são verificadas as leis bibliométricas – Lotka (1926), Bradford (1934) e Zipf (1949). A realização da análise bibliométrica ocorre via auxílio dos softwares Rstudio e Biblioshiny. Por sua vez, a revisão sistemática busca identificar lacunas de conhecimento na literatura existente. Essa parte da pesquisa ainda se encontra em fase de desenvolvimento. Os resultados parciais da pesquisa – decorrentes da análise bibliométrica – indicam que as palavras-chave com maior destaque são mercado de ação artificial, programação genética e modelagem baseada em agentes. Shu-Heng Chen aparece como autor e coautor mais relevante. Ao considerar as instituições às quais referidos autores estão vinculados, verifica-se que o país com maior quantidade de artigos publicados é a China. Entretanto, os artigos com maior número de citações estão vinculados a autores de instituições dos Estados Unidos. Há confirmação das três leis bibliométricas.

Palavras-chave: Mercados de ações artificiais, Modelagem baseada em agentes, Programação genética, Análise bibliométrica, Revisão sistemática

1. Introdução

Os mercados de ações artificiais (MAAs) são modelos de mercados reais, projetados com o objetivo de estudar sua dinâmica (Benhammada & Chikhi, 2010). Esses experimentos computacionais são capazes de capturar propriedades essenciais da realidade, a fim de reproduzir, analisar ou entender a sua dinâmica via interação de agentes. De um modo geral, a finalidade do MAA é reproduzir as características estatísticas do processo de preços - com hipóteses mínimas sobre a inteligência dos agentes (Ponta *et al.*, 2018). A interação de múltiplos agentes autônomos e inteligentes - como a origem de ambientes altamente dinâmicos e não determinísticos - contribui para a complexidade em aplicações como automação de mercados financeiros.

O MAA origina-se no sistema adaptativo complexo de Holland (1992). O sistema complexo é um conjunto de unidades de interação, cujas relações produzem propriedades que não são as das unidades individuais. Além disso, um sistema complexo é adaptativo, caso ele inclua unidades reativas. Essas são unidades capazes de exibir comportamentos diferentes em reação a condições ambientais alteradas (Tesfatsion, 2016). Ainda para Holland (1992), o sistema financeiro é um sistema interativo e dinâmico que muda em uma variedade de escalas de tempo individuais e evolutivas. Ele é composto por diversos agentes heterogêneos e interdependentes, possuidores de informações e decisões pessoais. Dado a isso, referido sistema possui a capacidade de mudar para se adequar ao ambiente, a partir da sua evolução e auto-organização.

Vários MAAs são encontrados na literatura, sendo simulados para prever ou compreender o preço (Yeh & Yang, 2010), risco (Chen & Huang, 2008) e retorno financeiro (Vanfossan *et al.*, 2020) das ações, assim como a perspectiva comportamental dos investidores (Rekik *et al.*, 2014). Um dos primeiros modelos a representar o mercado de ações é aquele criado pelo Instituto Santa Fé, descrito nas publicações de Palmer *et al.* (1994), Arthur *et al.* (1997) e LeBaron *et al.* (1999). Referido modelo analisa a interação entre agentes inteligentes, a fim de compreender os conceitos econômétricos ou fatos estilizados das séries temporais financeiras.

Na literatura, há trabalhos científicos que apresentam variados modelos para simulação dos MAAs. Porém, poucos deles se concentraram na revisão sistemática da literatura. A maioria dos artigos existentes apresenta um escopo limitado a temas específicos - análise da tecnologia computacional ou relacionados ao mercado financeiro (Wamba *et al.*, 2021; Ahmed *et al.*, 2022; Patel *et al.*, 2022). Assim, o principal objetivo deste estudo é realizar uma análise bibliométrica e revisão sistemática das publicações sobre o MAAs.

Para tanto, considera-se uma amostra final de 52 artigos, obtidos entre 01 de janeiro de 1.999 e 31 de julho de 2022. Esses artigos são extraídos das bases Web of Science (WoS) e Scopus. A análise bibliométrica descreve a frequência, citações e relações entre os principais dados dos artigos – autores, periódicos, palavras-chave etc. Além disso, são verificadas as leis bibliométricas – Lotka (1926), Bradford (1934) e Zipf (1949). Já a revisão sistemática oferece uma agenda para pesquisas futuras, a partir das lacunas de conhecimento verificadas na literatura existente. A realização de tais análises ocorre por meio dos softwares Rstudio e Biblioshiny.

2. Revisão da literatura

O mercado de ações é formado por investidores individuais e institucionais, cujas movimentações são impulsionadas por seu desejo de lucrar e aversão ao risco. Normalmente, as forças que motivam os investidores são ganância e medo. Além disso, as principais interações entre eles se dão por meio de troca de informações e transações. Assim, a tomada de decisão pelos indivíduos é feita de diferentes formas. Quando existe o risco de ganhos e

perdas, há a possibilidade de os investidores apresentarem um comportamento irracional. Quando se trata de ganhos, a maioria é avessa ao risco. Porém, quando se trata de perdas, os investidores são propensos a ele (Kahneman & Tversky, 1979).

Já para LeBaron *et al.* (1999), o mercado de ações é constituído por um grupo de agentes que interage entre si, promovendo sua adaptação às novas informações e atualização de seus conhecimentos. O desenvolvimento de um MAA - via experimentos computacionais - se dá por meio de uma coleção de procedimentos transformados em métodos. As unidades do modelo são denominadas de agentes. Cada agente é um *software* encapsulado que inclui dados e métodos comportamentais. Assim, os dados do agente, por exemplo, podem ser seus atributos, enquanto que os métodos comportamentais do agente podem ser suas estratégias ou seu algoritmo para atualização de estratégias. Esse encapsulamento em agentes tenta obter uma representação mais realista dos sistemas do mundo real.

Os MAAs distinguem-se dos modelos convencionais a partir de seus fundamentos. O modelador deve especificar seu estado e dados iniciais, definindo os comportamentos de cada agente. Esta deve ser a última interferência, uma vez que o modelo é dinamicamente completo. Significa dizer que o processo simulado deve ser capaz de se desenvolver ao longo do tempo - com base nas interações dos agentes - sem maiores intervenções do modelador (Chen & Yeh, 2002). Por sua vez, Ducha *et al.* (2021) entendem que os modelos baseados em agentes têm o objetivo de estudar diferentes aspectos da dinâmica do mercado. Influências aleatórias são comumente combinadas com dinâmicas endógenas, a fim de compreender como insumos - imprevisíveis do mundo exterior - originam a complexidade do comportamento do mercado.

Ainda sobre esse tema, Kwapień e Drożdż (2018) definem sistemas complexos como naturais ou sociais, sendo compostos por um grande número de elementos que interagem entre si. Tais sistemas exibem comportamento coletivo e - devido à troca de energia ou informação com o meio ambiente - eles podem facilmente modificar sua estrutura interna e padrões de atividade no seu processo de auto-organização. Como resultado, eles são flexíveis e se adaptam facilmente às condições externas variáveis. Entretanto, sua propriedade mais marcante é a capacidade de modelar fenômenos emergentes - que não podem ser previstos apenas a partir do conhecimento da estrutura dos sistemas e das interações entre seus elementos individuais.

Esta propriedade aponta para as abordagens holísticas que requerem descrições paralelas do mesmo sistema - em diferentes níveis da organização. Há evidências - igualmente consolidadas neste estudo - de que sistemas complexos diferentes - mesmo aparentemente distintos - podem ter características semelhantes, tanto em sua estrutura, quanto em seu comportamento. Assim, espera-se a existência de leis comuns e universais que rejam suas propriedades (Kwapień & Drożdż, 2018).

De fato, as premissas de homogeneidade e racionalidade perfeitas das finanças tradicionais enfrentam limitações teóricas e desafios empíricos. De acordo com He *et al.* (2019), empiricamente, os mercados financeiros apresentam características comuns - os denominados fatos estilizados e anomalias de mercado. Todavia, eles são difíceis de serem acomodados e explicados a partir do paradigma padrão das expectativas racionais. Essas anomalias podem ser caracterizadas pela ocorrência de fenômenos controversos no mercado. Entre eles, destacam-se: a) clusters persistentes de alta e baixa volatilidade, b) impulso no curto prazo e reversão no longo prazo, c) bolhas e quedas de preços e d) questionamentos sobre a Hipótese de Mercado Eficiente (HME).

A HME de Fama (1970) afirma que os mercados financeiros são eficientes em relação à informação. Assim, um investidor não é capaz de obter consistentemente retornos acima da média do mercado - com certo nível de risco - considerando apenas as informações

publicamente disponíveis no momento. Existem três formas de eficiência informacional, sendo: a) fraca - considera que os preços refletem todas as informações históricas publicamente disponíveis; b) semiforte - os preços refletem todas as informações históricas e atuais publicamente disponíveis; c) forte - os preços refletem a informação oculta ou privilegiada de *insiders* – não caracterizando eficiência de mercado.

Quando um agente econômico possui mais informação que outro, caracteriza-se a existência de assimetria informacional entre as partes (Akerlof, 1970). Em oposição à HME - que sustenta a eficiência em suas formas fraca ou semiforte - os preços de mercado dos ativos financeiros podem se desviar dos seus valores verdadeiros, devido a informações não reveladas. Ryu *et al.* (2021) esclarecem que há eventos de *disclosure* – como a emissão de relatórios de analistas e anúncios de ganhos – que mitigam a assimetria informacional nos mercados. Porém, isso não impede a existência de *insiders* que – por terem acesso a notícias privilegiadas - têm incentivos para explorar a superioridade da informação, enviando ordens antes de a notícia ser divulgada a todo o mercado.

Por sua vez, as preferências de um indivíduo são representadas por uma função de utilidade que especifica a vantagem obtida de cada resultado. Assim, a escolha de aquisição de uma ação pelo indivíduo produz a maior utilidade esperada, havendo um alinhamento entre suas preferências e probabilidades de ocorrência de resultados. Em suma, a determinação do valor de um item não pode ser baseada em seu preço, mas sim na utilidade que ele fornece. O preço de um item depende somente dele próprio, sendo para todos. Porém, sua utilidade depende das circunstâncias particulares do indivíduo que faz a estimativa. Embora essa estrutura seja fundamentada em axiomas razoáveis anteriores - e analiticamente tratáveis - sua validade descritiva é questionada, o que leva ao desenvolvimento de alternativas à utilidade esperada (Allais, 1953; Bernoulli, 1954; Kahneman & Tversky, 1979).

Quanto à teoria do prospecto, ela descreve um processo de pensamento intuitivo relacionado às formas como as pessoas avaliam o risco - ao tomar decisões de investimento. Suas principais premissas são: a) a preferência depende da referência - o valor de um resultado é dado pelo ganho/perda que ele representa sobre um ponto de referência, b) os indivíduos são avessos ao risco de perda - uma perda de uma determinada magnitude pesa mais que um ganho de mesma magnitude e c) ponderação de probabilidade - os indivíduos ponderam a probabilidade em decorrência de cada resultado.

Desde o artigo seminal de Kahneman e Tversky (1979) e Tversky e Kahneman (1992), a teoria do prospecto tem sido considerada uma das teorias descritivas mais relevante para retratar o comportamento dos indivíduos tomadores de decisão. Investidores individuais são percebidos como menos racionais. Logo, suas decisões de investimento são mais propensas ao viés cognitivo, desviando-se das previsões da teoria da utilidade esperada (Wang *et al.*, 2021).

3. Metodologia

O objetivo deste estudo é realizar uma análise bibliométrica e revisão sistemática das publicações sobre os MAAs. Para tanto, são implementadas 7 etapas. As etapas 1 a 5 estão relacionadas à análise bibliométrica, enquanto que nas etapas 6 e 7 é desenvolvida a revisão sistemática.

Etapa 1 – Escolha da base de dados. Os artigos são obtidos a partir das bases de dados: (i) Web of Science (WoS) - pertencente à Clarivate Analytics. A WoS contém registro de artigos científicos publicados em periódicos com alto fator de impacto, sendo classificados por meio do índice Journal Citation Reports (JCR); e (ii) Scopus – base de dados de artigos

acadêmicos da Elsevier. A Scopus mensura a relevância dos periódicos acadêmicos por meio do índice CiteScore.

Etapa 2 – Definição dos parâmetros iniciais da pesquisa. Para limitar o alcance do estudo, é realizada uma pesquisa por meio de palavras-chave. Na sequência são aplicados filtros das próprias bases da WoS e Scopus relacionados ao idioma, tipo de documento, área, categoria e período.

Etapa 3 – Unificação das bases de dados e exclusão de artigos em duplicidade. Após o download dos arquivos - das bases da WoS e Scopus – ocorre sua unificação. Sua padronização para fins de análise via *software* Biblioshiny é estabelecida com base nos campos da WoS. Na sequência, há a exclusão dos artigos em duplicidade. Assim, mantém-se apenas um único documento, evitando sua repetição na amostra intermediária. Ambas ações são realizadas via *software* R Studio.

Etapa 4 – Exclusão de artigos não relacionados ao tema definido. Uma vez definida a amostra intermediária, efetua-se a leitura do resumo, introdução e conclusão desses artigos. A amostra final, então, é obtida a partir da exclusão daqueles que não estejam alinhados com o tema principal do estudo. A evolução da amostra final está apresentada na Tabela 1.

Etapa 5 – Análise bibliométrica. O arquivo da amostra final é importado para o *software* Biblioshiny, a partir do qual é efetuada a análise dos dados objetivos dos artigos – países, autores, palavras-chave, instituições etc. - para elaboração e análise de tabelas e mapas de relacionamento/cocitação. Adicionalmente, as análises realizadas são complementadas pela verificação das principais leis bibliométricas, sendo: a) Lei de Zipf (1949) - categorização e estimativa da frequência das palavras-chave com o auxílio do *software* Rank Words; b) Lei de Bradford (1934) – verificação dos periódicos; e c) Lei de Lotka (1926) – identificação dos pesquisadores que possuem maior frequência de produção no tema definido.

Tabela 1 - Evolução da amostra final

Sinal	Descrição	Número de artigos	
		WoS	Scopus
(+)	Amostra inicial , obtida por meio das palavras-chave: “artificial stock market”.	176	241
(-)	Idiomas diferentes de “english”	2	16
(-)	Tipos de documentos diferentes de “article” e “review”	69	103
(-)	Áreas de pesquisa diferentes de: a) WoS: Business finance b) Scopus: i) Economics, econometrics and finance, ii) Business, management and accounting, iii) Decision sciences	71	70
(=)	Subtotal	34	52
(-)	Exclusão de artigos em duplicidade nas bases da WoS e Scopus	29	
(=)	Amostra intermediária total	57	
(-)	Artigos para os quais não foi possível obter o arquivo com o texto completo ou converter seus dados para análise da Lei de Zipf (1949) pelo <i>software</i> Rank Words	5	
(=)	Amostra final total	52	

Etapa 6 – Leitura e codificação dos artigos. Identificação dos objetivos, amostras, métodos, contribuições e demais características dos artigos. Estes são classificados e codificados em categorias e subcategorias não exclusivas. Isso significa que um mesmo artigo pode ser classificado em mais de uma subcategoria, o que possibilita que a soma da frequência das classificações seja superior a 100%.

Etapa 7 – Revisão sistemática. Realização da contagem de frequência das subcategorias – e análise de suas combinações menos frequentes - de forma a possibilitar a identificação de lacunas do conhecimento.

Tabela 2 - Matriz de (sub) categorização

#	Categorias	#	Subcategorias	Definição/Conceito
1	Tema principal	1.1	Mercados de ações	Ambiente em que empresas de capital aberto negociam frações de seu capital social.
		1.2	Mercados de ações artificiais	Modelo baseado em sistemas computacionais que simulam o ambiente de um mercado de ações real, tendo como objetivo estudar sua dinâmica
		1.3	Sistemas complexos	Sistema com componentes que não possuem um controle central e atuam de acordo com regras simples de operação, dando origem a comportamentos coletivos complexos, processamento de informação sofisticado e adaptação via aprendizagem ou evolução.
		1.4	Fatos estilizados	Aproximação teórica de um fenômeno observado empiricamente
		1.5	Outros	Demais temas não relacionados às subcategorias 1.1 a 1.4
2	Teorias relacionadas	2.1	Hipótese do mercado eficiente	Um mercado é eficiente quando seus participantes precisam ativos baseados em toda a informação disponível sobre eventos que possam influenciar o retorno do ativo negociado no mercado. Há 3 formas de eficiência informacional - fraca, semi-forte e forte. A eficiência de mercado ocorre somente nas formas fraca e semi-forte (Fama, 1970, 1991)
		2.2	Assimetria informacional	Descrição de um fenômeno segundo o qual alguns agentes econômicos têm mais informações do que outros (Akerlof, 1970)
		2.3	Teoria do prospecto	Trata-se de uma crítica ao modelo da utilidade esperada como uma forma de tomada de decisão pelos indivíduos. Os indivíduos fazem escolhas diferentes quando existe o risco de ganhos e perdas, além da falta total de racionalidade. Quando se trata de ganhos, a maioria é avessa ao risco. Porém, quando se trata de perdas, elas são propensas a ele (Kahneman & Tversky, 1979).
		2.4	Utilidade esperada	A determinação do valor de um item não pode ser baseado em seu preço, mas sim na utilidade que ele fornece. O preço de um item depende somente do próprio item e é igual para todo mundo. Porém, a utilidade de um item depende das circunstâncias particulares do indivíduo que faz a estimativa (Bernoulli, 1954)
		2.5	Outros	Demais teorias não relacionadas às subcategorias 2.1 a 2.4
3	Tipos de simulações computacionais	3.1	Modelos baseados em agentes	Modelos computacionais que simulam a especificação das ações por meio de interações entre agentes autônomos. Eles são mais intuitivos que os modelos matemáticos ou estatísticos, uma vez que representam objetos mais próximos à realidade - elementos individuais no mundo.
		3.2	Modelos autômatos celulares	Modelos de evolução temporal mais simples; porém, com capacidade para apresentar comportamentos mais complexos. Eles consistem num conjunto de células, no qual cada uma delas assume um estado em um sistema finito de possibilidades.
		3.3	Programação evolucionária	Programação baseada no desenvolvimento de algoritmos evolucionários. Tais algoritmos são capazes de encontrar soluções adequadas para problemas complexos. Eles utilizam princípios básicos de teoria da evolução e genética, podendo

			ser modelados com poucas e simples linhas de código. Os algoritmos genéticos também são utilizados no desenvolvimento de inteligência artificial
3	3.4	Sistemas dinâmicos	Sistemas no qual pelo menos uma de suas variáveis de estado depende do tempo. Sua regra de evolução é definida por uma função que descreve quais estados futuros seguem a partir do estado atual. Assim, eles realizam simulações via retroalimentação, afetando o comportamento do sistema como um todo.
		Outros	Demais tipos de simulação não relacionados às subcategorias 3.1 a 3.4
4	4.1	Tipos de investidores	Há 3 principais tipos - conservador, moderado e arrojado ou agressivo. Cada um deles possui uma finalidade em suas operações, atuando como hedger, arbitrador ou especulador
		Mecanismos de investimento	Como os investidores são inseridos no mercado.
		Processos de tomada de decisão	Seleção e escolha de uma opção, a partir da exclusão de outras alternativas possíveis
		Estratégias de liquidação	A liquidação está relacionada à retirada da participação dos investidores do capital próprio da empresa. Essa decisão está relacionada com estratégias de maximização de lucro ou o objetivo pré-estabelecido.
		Outros	Demais estruturas de mercado não relacionados às subcategorias 4.1 a 4.4
5	5.1	Preço da ação	Definição do critério de precificação inicial da ação no modelo
		Tipos de negociações	Definição dos tipos de negociações entre investidores e instituições previstas pelo modelo
		Quantidade de investidores	Número de investidores considerados pelo modelo
		Instituições reguladoras	Definição sobre a inclusão ou não de instituições reguladoras no modelo
		Aprendizagem dos investidores	Definição sobre a existência do processo de aprendizagem pelos investidores no modelo. Refere-se à capacidade de os investidores ampliarem e atualizarem seu conhecimento - a partir das informações geradas nos pregões anteriores - em sua próxima decisão
		Outros	Demais parâmetros não relacionados às subcategorias 5.1 a 5.5
6	6.1	De 1 a 50 pregões	O modelo considera o histórico de preço/retorno das ações de 1 a 50 pregões
		De 51 a 100 pregões	O modelo considera o histórico de preço/retorno das ações de 51 a 100 pregões
		De 101 a 200 pregões	O modelo considera o histórico de preço/retorno das ações de 101 a 200 pregões
		De 201 a 300 pregões	O modelo considera o histórico de preço/retorno das ações de 201 a 300 pregões
		De 301 a 500 pregões	O modelo considera o histórico de preço/retorno das ações de 301 a 500 pregões
		Outros	Demais intervalos de pregões não relacionados às subcategorias 6.1 a 6.5
7	7.1	Estados Unidos	O modelo considera dados relacionados ao mercado de ações norte-americano
		Europa	O modelo considera dados relacionados ao mercado de ações europeu
		Ásia	O modelo considera dados relacionados ao mercado de ações asiático

		7.4	Outros	Demais países não relacionados às subcategorias 7.1 a 7.3
		7.5	Sem comparação com países	O modelo não considera dados do mercado de qualquer país
8	Análise do modelo	8.1	Calibração do modelo	Calibração obtida via mudança de parâmetros capazes de ajustar como o modelo opera e simula o processo. Tais parâmetros influenciam o comportamento do agente.
		8.2	Verificação do modelo	A verificação se dá pela análise dos dados de saída do modelo. Eles são comparados com o que é esperado, a partir dos dados de entrada desse mesmo modelo
		8.3	Validação do modelo	A validação ocorre via comparação dos resultados com os valores esperados - sendo aqueles identificados a partir dos dados históricos da área de estudo. É esperado que o modelo produza resultados semelhantes aos historicamente verificados.
		8.4	Análise de sensibilidade	Exame do efeito da variação - de um determinado item - em relação ao seu valor original.
		8.5	Outros	Demais análises não relacionadas às subcategorias 8.1 a 8.4
9	Tipos de modelos	9.1	Invariante no tempo	Verificação se o sistema é invariante no tempo. Caso seja, suas características não devem alterar com o tempo
		9.2	Estáticos ou dinâmicos	Modelos estáticos representam o estado de um sistema em um determinado instante. Em sua definição, a variável tempo não é considerada. Já os modelos dinâmicos são formulados para representar alterações de estado do sistema ao longo do tempo de simulação
		9.3	Determinísticos ou estocásticos	Modelos determinísticos não fazem uso de variáveis aleatórias. Ao contrário, modelos estocásticos não oferecem soluções únicas, mas apresentam uma distribuição de soluções associadas a uma probabilidade.
		9.4	Tempo real ou simulado	Nesse modelo, a escala de tempo é a real. Os eventos ocorrem e são tratados na mesma escala de tempo correspondente ao sistema real.
		9.5	Discreto ou contínuo	Modelos discretos são aqueles em que o avanço da contagem de tempo na simulação se dá na forma incremental. Os valores podem ser definidos em função da ocorrência dos eventos ou pela determinação de um valor fixo. Já para os modelos contínuos, o avanço da contagem de tempo na simulação dá-se de forma contínua. Isso possibilita determinar os valores das variáveis de estado a qualquer instante.
		9.6	Outros	Demais tipos de modelos não relacionados às subcategorias 9.1 a 9.5
10	Análise dos resultados	10.1	Testes estatísticos clássicos	Testes que determinam se dois conjuntos de dados podem ser considerados semelhantes ou não - ex: teste t e qui-quadrado
		10.2	Abordagem por inspeção	Nessa abordagem, são utilizadas medidas de estatística descritiva de posição central, dispersão e relação dos dados - ex: média, variância, correlação e histogramas
		10.3	Abordagem do intervalo de confiança	O intervalo de confiança padrão é considerado ao nível de 95%. Entretanto, quando apresentado graficamente, ele pode apresentar outros níveis como 90% e 99%. Certos fatores podem afetar o tamanho do intervalo de confiança, incluindo tamanho da amostra, nível de confiança e variabilidade da população.
		10.4	Outros	Demais análises não relacionadas às subcategorias 10.1 a 10.3
11	Fatos estilizados de retorno e	11.1	Não normalidade da distribuição dos retornos	Distribuição leptocúrtica dos dados, sendo caracterizada pela existência de caudas gordas - curtose (k) > 0

	de volatilidade	11.2	Ausência de autocorrelação linear dos retornos	A autocorrelação é uma medida que informa o quanto o valor anterior de uma variável aleatória é capaz de influenciar seu próximo valor. A autocorrelação de funções não lineares é geralmente positiva e tende a decair em um ritmo muito mais lento. Para uma melhor adequação do modelo, é desejável que a autocorrelação seja inexistente
		11.3	Agrupamentos de volatilidade	A volatilidade é a quantidade e intensidade de flutuações e oscilações que ocorrem com uma série de retornos. Agrupamentos de volatilidade são séries temporais caracterizadas por grandes e intermitentes flutuações
		11.4	Efeito de alavancagem	O efeito de alavancagem explica a rentabilidade de uma empresa em termos do retorno sobre o capital empregado e do custo da dívida. A maioria das medidas de volatilidade de um ativo é negativamente correlacionada com os seus preços passados. Desta forma, em geral a volatilidade é maior em momentos de baixa de seus papéis, possivelmente devido ao aumento subsequente da alavancagem da companhia nestes períodos. Assim, a volatilidade assimétrica captura essa tendência assimétrica da volatilidade de ser maior durante quedas de preços do que durante aumentos de preços das ações
		11.5	Correlação entre volume e volatilidade	A correlação retorno-volume apresenta o mesmo comportamento que a alavancagem. Isso significa que quedas de retornos estão associadas a um aumento na atividade de negociação
		11.6	Outros	Demais fatos estilizados não relacionados às subcategorias 11.1 a 11.5
12	Conclusão	12.1	Assertividade do modelo	Os resultados são consistentes com as métricas utilizadas.
		12.2	Não assertividade modelo	Os resultados não são consistentes com as métricas utilizadas
		12.3	Outros	Demais conclusões não relacionadas às subcategorias 12.1 a 12.2
13	Avenidas para o futuro	13.1	Calibração de dados	Os modelos são sensíveis aos dados empíricos. Logo, sua calibração deve possibilitar maior confiabilidade aos modelos
		13.2	Variação de ativos	Os títulos devem variar conforme as estratégias de liquidação das carteiras.
		13.3	Ampliação das estratégias de investimentos	Ampliação das sugestões de investimento aos agentes, bem como do número de interações que compõem cada tentativa de investimento
		13.4	Foco na microestrutura de mercado	Entendimento da microestrutura e dos regulamentos do mercado financeiro, a fim de que os comerciantes possam obter mais lucro
		13.4	Outros	Outras propostas de pesquisas futuras não relacionadas às subcategorias 13.1 a 13.4

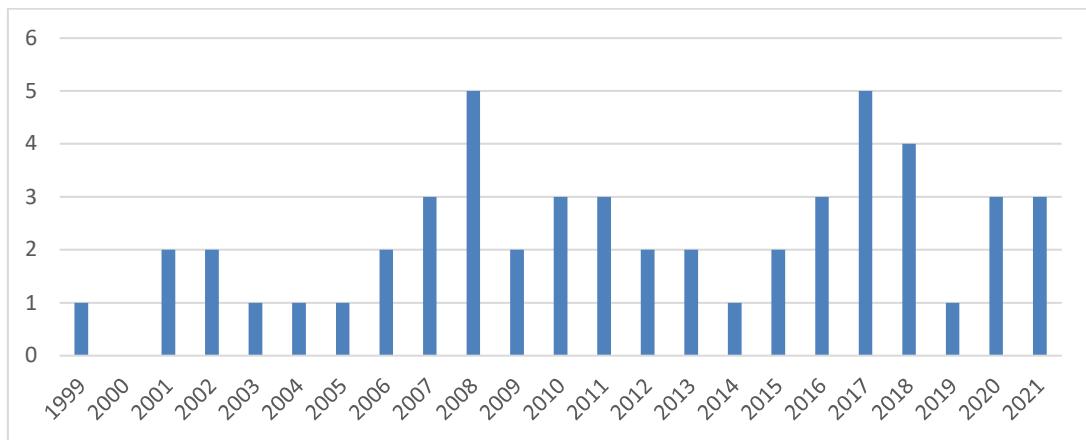
4. Análise de resultados

O item 4.1 apresenta os resultados preliminares da análise bibliométrica, mencionados na Etapa 5 da Metodologia. Por sua vez, o item 4.2 conterá os resultados da revisão sistemática - a ser posteriormente desenvolvida - cujas etapas estão descritas nas Etapas 6 e 7 do item 3 deste estudo.

4.1. Análise bibliométrica

A amostra final é composta por 52 artigos, obtidos entre 01 de janeiro de 1999 e 31 de julho de 2022, a partir das bases da WoS e Scopus – vide Tabela 1. A Figura 1 indica que há de 1 a 5 artigos publicados em cada ano, havendo destaque para os anos de 2008 e 2017. O artigo pioneiro sobre MAA é o LeBaron *et al.* (1999). Nele, o mercado é modelado por algoritmos de inteligência artificial que assumem o papel de *traders*. Além disso, o ambiente simulado é capaz de replicar vários fenômenos, incluindo previsibilidade fundamental e técnica, persistência de volatilidade e leptocurtose.

Figura 1 - Distribuição dos artigos no tempo



Fonte: Autores

Por sua vez, a Tabela 3 apresenta a frequência das 10 principais palavras-chave da amostra final - de um total de 279, destacando-se: MAA (5,73%), programação genética (5,38%) e modelagem baseada em agentes (4,66%). Em síntese, os artigos da amostra discutem, sobretudo, tema associados ao MAA e aos métodos utilizados para a sua construção, destacando-se: programação genética e modelagem baseada em agentes. Referidas palavras podem ser igualmente observadas na Figura 2.

Tabela 3 – Distribuição das palavras-chave

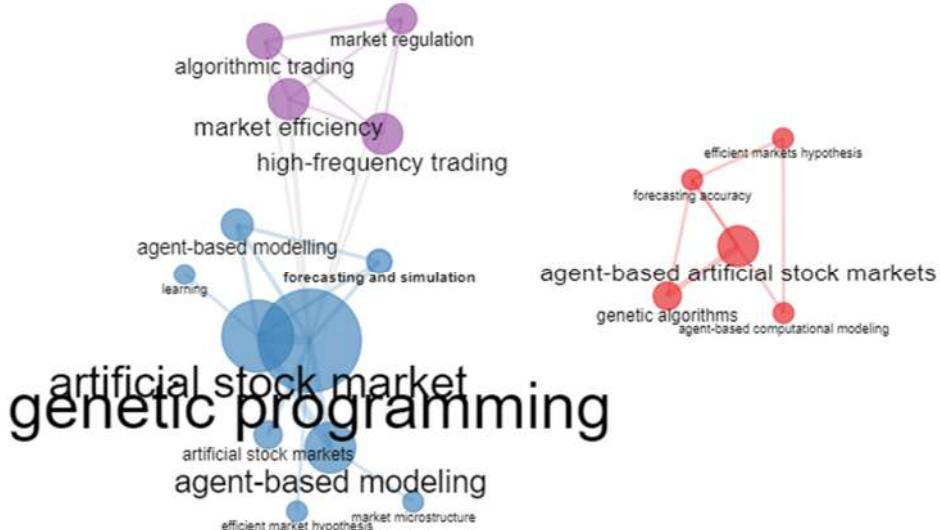
Palavras chave	Ocorrências	%
Mercado de ação artificial	16	5,73%
Programação genética	15	5,38%
Modelagem baseada em agentes	13	4,66%
Mercados de ações artificial baseados em agente	7	2,51%
Algoritmos genéticos	5	1,79%
Negociação de alta frequência	5	1,79%
Eficiência de mercado	5	1,79%
Modelo baseado em agente	4	1,43%
Negociação algorítmica	4	1,43%
Aprendizagem	4	1,43%

Fonte: Biblioshiny

A Figura 2 exibe o mapa de coocorrências das palavras-chave mais utilizadas nos artigos. Novamente, destacam-se as palavras programação genética, MAA, modelagem baseada em agentes e mercados de ações artificiais baseados em agentes. A relação entre

MAA e programação genética indica que - em alguns modelos - são utilizadas linguagens que apresentam a evolução dos programas, sendo semelhante à evolução biológica natural.

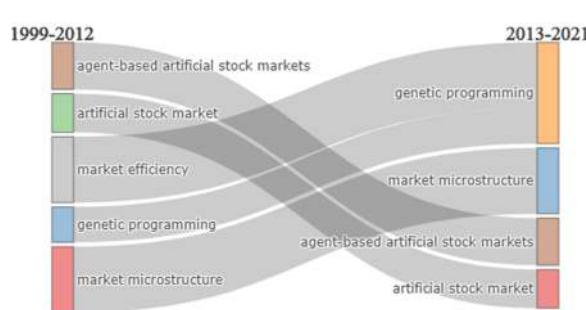
Figura 2 - Mapa de coocorrência entre as palavras-chave mais citadas



Fonte: Biblioshiny

Na Figura 3, observa-se a evolução das palavras-chave durante os anos 1999 a 2021. Há uma mudança das palavras de eficiência de mercado para programação genética, refletindo a importância da programação para esse conceito de mercado. Já na Figura 4, confirma-se a relevância das três principais palavras-chave - MAA, programação genética e modelagem baseada em agentes - por meio do seu número de ocorrências, à semelhança das Tabela 3 e Figuras 2 e 3.

Figura 3 - Mapa da evolução temática das palavras-chave



Fonte: Biblioshiny

Figura 4 - Nuvem das palavras-chave



Fonte: Biblioshiny

Ainda sobre as palavras-chave, a Lei de Zipf (1949) analisa e quantifica a distribuição das palavras num determinado texto. Para isso, considera-se o princípio de que há uma tendência de uso mínimo das palavras que possuem alta frequência de ocorrência. A primeira lei de Zipf aponta que (r) é a ordem da série, (f) a frequência de ocorrência e (C) a constante para qualquer texto – vide Equação 1:

$$r \times f = C \quad (1)$$

Por sua vez, para as palavras com baixa frequência, Zipf propôs uma segunda lei, modificada e revisada por Booth (1967). Para o autor, em um determinado texto, várias palavras com baixa frequência de ocorrência possuem a mesma assiduidade – vide Equação 2:

$$I_n = 2I_1 / n(n+1) \quad (2)$$

Em que:

I_1 = número de palavras que têm frequência 1

I_n = número de palavras que têm frequência n

n = ponto de Goffman ou local de transição das palavras de baixa para as de alta frequência

As leis de Zipf definem as extremidades da lista de distribuição de palavras de determinado texto. Deste modo, entre esses pontos extremos, há uma região de transição das palavras de alta para as de baixa frequência. Para Goffman (1971), a existência de um ponto de transição entre as palavras de maior e menor frequência representam mais adequadamente o conteúdo semântico de um determinado texto. Pao (1978) apresenta a fórmula do ponto transição de Goffman – vide Equação 3.

$$T = (-1 + \sqrt{1 + 8I_1}) / 2 \quad (3)$$

Em que:

T = ponto de transição de Goffman

I_1 = número de palavras que têm frequência 1

A identificação do ponto T de Goffman ocorre por meio do ordenamento decrescente das palavras no *software Rank Words*. Na sequência, são identificadas aquelas que se repetem apenas uma única vez para cálculo do ponto T de Goffman. Em seguida, são localizadas as palavras que estão acima da classificação indicada por esse ponto. A Tabela 4 apresenta os 10 primeiros artigos da amostra final - por ordem decrescente do ponto T de Goffman. Nela, o ponto de transição das palavras varia entre 43,32 e 60,03, sendo que sua média é de 48,64.

No caso do artigo de Westerhoff (2006), por exemplo, há 1.832 palavras cuja frequência de repetição é igual a 1 - ex: *specialist, results, produces* etc. O cálculo da Equação 3 resulta no valor de 60,03 $(-1 + \sqrt{1 + 8 * 1.832}) / 2$ pontos do T de Goffman. A palavra que mais se aproxima dessa frequência é *agent*. O autor propõe um MAA com dois tipos de *traders* - fundamentalistas e grafistas. O modelo é capaz de replicar vários fatos estilizados e o resultado indica que os intervalos de negociação têm o poder de reduzir a volatilidade.

Tabela 4 - T de Goffman

No.	Autor	T de Goffman
1	Westerhoff (2006)	60,03
2	Oldham (2020)	56,01
3	Chen & Tokinaga (2004)	54,18
4	Benink (2010)	51,07
5	Mandes (2020)	45,19
6	Manahov (2016)	44,46

7	Schmitt & Westerhoff (2017)	44,42
8	Chen & Huang (2007)	44,02
9	Tsao & Huang (2018)	43,68
10	Mandes & Winker (2017)	43,32

Fonte: Autores

Na sequência, realiza-se uma análise da região em que estão as palavras de maior frequência relacionadas à temática principal – para cada um dos 52 artigos da amostra final. O Rank Words classifica as palavras por ordem decrescente de repetição. São excluídas aquelas não relevantes ao estudo – ex: artigos (in)definidos, preposições, advérbios e pronomes. Das palavras restantes, identifica-se aquelas com maior frequência.

A Tabela 5 apresenta os 10 primeiros artigos por ordem decrescente da palavra com maior repetição - em cada estudo. Por exemplo, Yang *et al.* (2020) descrevem um MAA - baseado em agentes - com o intuito de analisar a mínima alternação no tamanho parcial de uma ação (*tick size*) na qualidade do mercado. Os resultados apresentam implicações práticas sobre as configurações de *tick size*, investigando sistemas de dimensionamento baseados em volume.

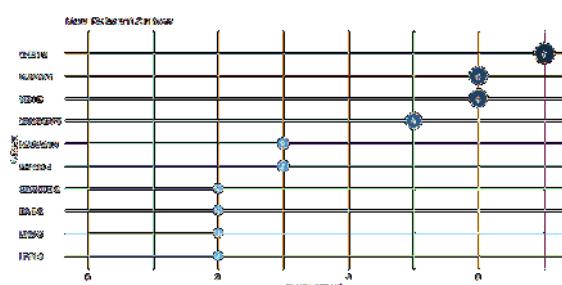
Tabela 5 - Lei de Zipf

No.	Autor	Palavra	Quantidade	Frequência
1	Yang <i>et al.</i> (2020)	<i>Tick</i>	308	12,67%
2	Yeh & Yang (2013)	<i>Price</i>	276	11,36%
3	Yeh (2007)	<i>Market</i>	266	10,95%
4	Zhang <i>et al.</i> (2018)	<i>Trading</i>	256	10,53%
5	Manavoh (2016)	<i>Trading</i>	251	10,33%
6	Manavoh (2016)	<i>Trading</i>	236	9,71%
7	Kluger & McBride (2011)	<i>Agents</i>	230	9,47%
8	Mandes (2020)	<i>Market</i>	224	9,22%
9	Raberto <i>et al.</i> (2003)	<i>Model</i>	192	7,90%
10	Manavoh (2016)	<i>Trading</i>	191	7,86%
Total			2.430	100,00%

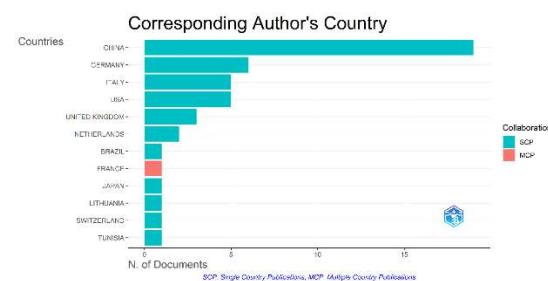
Fonte: Autores

Com relação à autoria dos artigos, são identificados 89 autores. Destes, 10 publicam individualmente e 79 são coautores. A Figura 5 apresenta os 10 autores mais relevantes em termos de número de publicações e o total de citações por ano ao longo do tempo. Shu-Heng Chen aparece como autor e coautor mais relevante, tendo participado de 7 artigos (13,5%). Neles, o autor aborda os seguintes temas: programação genética, modelagem baseada em agentes, teste de causalidade de Granger, hipótese de mercado eficiente e agentes autônomos.

Para os 12 países da amostra, a Figura 6, indica a proporção entre artigos escritos por autores associados a instituições de um mesmo país - *single country publications* (SCP) e aqueles escritos por autores associados a instituições de diferentes países - *multiple country publications* (MCP). Os autores que mais publicam são aqueles vinculados a instituições localizadas na China. Nelas, de um total de 19 artigos, todos são publicações intra-país. Num segundo plano, estão Alemanha e Itália. Nesses países, verifica-se a publicação de 6 e 5 artigos, respectivamente, sendo que todos são SCP. Somente a França apresenta 1 artigo com colaboração de autores vinculados a instituições de outros países - MCP.

Figura 5 - Autores mais relevantes

Fonte: Biblioshiny

Figura 6 – Colaboração entre países na publicação de artigos

Fonte: Biblioshiny

Por sua vez, a Tabela 6 analisa as citações dos 52 artigos da amostra final. Do total de 1.240, 505(40,73%) citações são de artigos elaborados por autores associados a instituições localizadas nos Estados Unidos (USA). Destaca-se ainda que a China – não obstante ao fato de ter mais autores – seus artigos têm menos citações – 475(38,31%) que os dos USA.

Tabela 6 - Total de citações

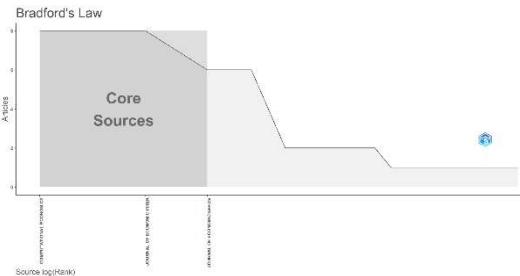
País	Total de citações	Frequência
USA	505	40,73%
China	475	38,31%
Italy	87	7,02%
Germany	83	6,69%
Brazil	45	3,63%
Lithuania	11	0,89%
Netherlands	9	0,73%
Tunisia	9	0,73%
United Kingdom	8	0,65%
France	5	0,40%
Japan	2	0,16%
Switzerland	1	0,08%
Total	1.240	100,0%

Quanto à Figura 7, ela apresenta a aplicação da Lei de Bradford (1934) sobre periódicos - há poucos periódicos que produzem muitos artigos e muitos periódicos que produzem poucos artigos sobre um determinado tema. Brookes (1969) cita que a Lei de Bradford estima o grau de relevância dos periódicos acadêmicos em áreas específicas do conhecimento. Assim, caso os periódicos sejam classificados por ordem decrescente de produtividade, eles podem ser distribuídos em zonas com variação na proporção $1: n: n^2$ e assim por diante. Referidas zonas são formadas ao dividir o total de artigos publicados por três.

Assim, tem-se que na zona 1 – de um total de 52 artigos e 23 periódicos, 3(13,04%) periódicos publicam 22(42,30%) artigos. Nela, há periódicos que publicam de 8 - Computational Economics e Journal of Economic Interaction and Coordination - a 6 - Journal of Economic Behavior and Organization - artigos. Por sua vez, na zona 2, há 5 (21,74%) periódicos que publicam 14 (26,92%) dos artigos. Em média esses periódicos têm 2,26 artigos publicados. Por fim, na zona 3, há 15(65,21%) periódicos que publicam 16(30,77%) artigos. Destaca-se que - nesta última zona - a média de artigos publicados por periódico é ainda menor, sendo de 1 artigo por *journal*.

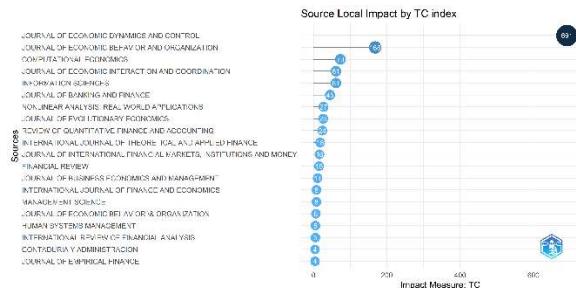
Ademais, a Figura 8 apresenta o total de citações (TC) obtidas por um determinado periódico. Referidas citações – também denominadas de *global citations* – significa o total de citações que um artigo - incluído na amostra final - tem recebido de outros artigos indexados nas bases de dados bibliográficas da WoS e Scopus. Para os 20 periódicos mais citados, há um total de 1.281 citações. Entre eles, destacam-se o Journal of Economic Dynamics and Control (691 ou 53,94%), Journal of Economic Behavior and Organization (168 ou 13,11%) e Computational Economics (73 ou 5,69%).

Figura 7 - Lei de Bradford sobre periódicos



Fonte: Biblioshiny

Figura 8 – Total de citações por periódico - TC index



Fonte: Biblioshiny

Já a Figura 9 apresenta as 10 principais instituições com mais de 2 artigos publicados. Assim, ela analisa 71,15% (37/52) dos artigos da amostra final. A National Chengchi Universityy - localizada em Taipei City/Taiwan - apresenta 13,46% (7/52) dos artigos publicados. Em seguida, destaca-se a Lunghwa University of Science and Technology - estabelecida na Taoyuan City/Taiwan – com 9,61% (5/52) dos artigos. Outras 8 universidades publicam 48,08% (25/52) dos artigos – cada.

Lotka (1926) afirma que um pequeno número de autores produz muitos artigos e que a produção obtida por esse reduzido número de pesquisadores se iguala em quantidade ao desempenho dos demais. Essa lei é denominada de lei do quadrado inverso - vide Equação 4.

$$a_n = a_1 / n^2, n = 1, 2, 3... \quad (4)$$

Em que:

a_n = número de autores que publicaram n artigos

a_1 = número de autores que publicaram um artigo

n = número de artigos publicados por autor

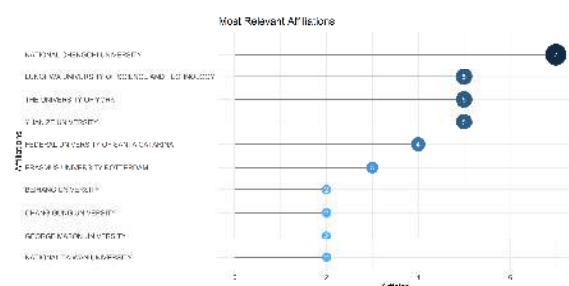
Na Equação 5, Chung e Cox (1990) esclarecem que o número de autores com um único artigo publicado - conforme a Lei de Lotka - seria de 60,8%. Ademais, o número de autores com dois artigos publicados deveria ter a frequência de 15,2% (60,8%/22). Já autor com três artigos publicados seria de 6,8% (60,8%/32).

$$a_1 = 6 / \pi^2 = 0,6079 = 60,8\% \quad (5)$$

A Figura 10 apresenta a quantidade e percentual de artigos publicados pelos autores da amostra. Os valores padrão da Lei de Lotka estão apresentados nas variações da Equação 5. Verifica-se que, de fato, há um percentual menor de autores publicando uma maior quantidade de artigos, confirmando a Lei de Lotka. 1 (1,1%) dos 89 autores – da amostra

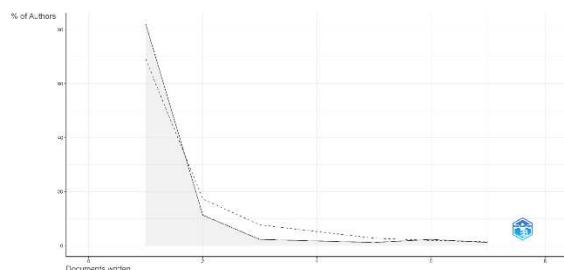
final - é responsável pela publicação de 7 artigos relacionados ao tema MAA. Já 15 (16,85%) autores publicam entre 2 e 6 artigos sobre o tema. Outros 73 (82%) publicam 1 único artigo relacionado a esse tema, o que caracteriza uma produtividade ainda menor que a referência padrão obtida por Lotka (60,8%).

Figura 9 – Principais filiações



Fonte: Biblioshiny

Figura 10 – Quantidade e percentual de artigos publicados – Lei de Lotka



Fonte: Biblioshiny

4.2 Revisão Sistemática

A ser desenvolvida

5. Conclusão

Este estudo tem como objetivo realizar uma análise bibliométrica e revisão sistemática das publicações sobre os MAAs. A amostra final é composta por 52 artigos, considerando o período entre 01 de janeiro de 1.999 e 31 de julho de 2022. Tais artigos são obtidos a partir das bases WoS e Scopus. A análise bibliométrica descreve a frequência, citações e relações entre os principais dados dos artigos – autores, periódicos, palavras-chave etc.

Além disso, são verificadas as leis bibliométricas – Lotka (1926), Bradford (1934) e Zipf (1949). A realização da análise bibliométrica ocorre via auxílio dos softwares Rstudio e Biblioshiny. Por sua vez, a revisão sistemática busca identificar lacunas de conhecimento na literatura existente. Referidas lacunas apontam para avenidas futuras – de temas relacionados - a outros pesquisadores. Essa parte da pesquisa ainda se encontra em fase de desenvolvimento.

Os resultados parciais da pesquisa – decorrentes da análise bibliométrica – indicam que as palavras-chave com maior destaque são mercado de ação artificial (5,73%), programação genética (5,38%) e modelagem baseada em agentes (4,66%) – vide Tabela 3. Entretanto, a palavra programação genética tem maior visibilidade entre 2013 e 2021 – vide Figura 3. De acordo com a lei de Zipf, as palavras com maior frequência são *tick, price, market, trading, agentes e model* – vide Tabela 5.

Sobre os autores dos artigos da amostra final, Shu-Heng Chen aparece como autor e coautor mais relevante, tendo participado de 7 dos 52 artigos (13,5%) da amostra final – vide Figura 5. Tal fato confirma a lei de Lotka. Outros 15 (16,8%) dos 89 autores publicam entre 2 e 6 artigos sobre o tema, enquanto que os demais 73 (82%) publicam um único artigo – vide Figura 10. Ao considerar as instituições às quais referidos autores estão vinculados, verifica-se que o país com maior quantidade de artigos publicados é a China – 19(36,5%) seguido da Alemanha e Itália.

Tais instituições estão sediadas em 12 países. Porém, entre eles, apenas a França possui artigos publicados com autores de outros países – vide Figura 6. Cabe destacar que, não obstante o fato de a China ter uma maior quantidade de artigos publicados, aqueles com maior número de citações – 505(40,73%) – estão vinculados a autores de instituições norte americanas – vide Tabela 6. Entre as instituições de ensino com maior volume de artigos publicados estão a National Chengchi Universityy – com 13,46% (7/52) – e Lunghwa University of Science and Technology - com 9,61% (5/52) – vide Figura 9.

Quanto aos periódicos, também é confirmada a lei de Bradford - de um total de 52 artigos e 23 periódicos, 3(13,04%) periódicos publicam 22(42,30%) artigos. Entre eles, destacam-se Computational Economics Journal of Economic Interaction and Coordination e Journal of Economic Behavior and Organization – vide Figura 7. Por sua vez, a relevância dos periódicos pode ser verificada por meio do número de citações que seus artigos recebem. De um total de 1.281 citações, aqueles que apresentam maior frequência são: Journal of Economic Dynamics and Control (691 ou 53,94%), Journal of Economic Behavior and Organization (168 ou 13,11%) e Computational Economics (73 ou 5,69%) – vide Figura 8.

Quanto às limitações deste estudo, os resultados apresentados referem-se ao conjunto de artigos dos periódicos que atendem aos critérios de seleção estabelecidos – vide Tabela 1 - e às bases de dados utilizadas – WoS e Scopus. Logo, para sua evolução, sugere-se a análise de artigos que considerem outras bases de artigos acadêmicos, bem como a investigação sobre os temas de programação genética e modelagem baseada em agentes.

Referências

- Ahmed, S., Alshater, M.M., Ammari, A., & Hammami, H. (2022). Artificial intelligence and machine learning in finance: A bibliometric review. *Research in International Business and Finance*, 61, 1-14. <https://doi.org/10.1016/j.ribaf.2022.101646>.
- Akerlof, G.A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3), 488-500. <http://www.jstor.org/stable/1879431>.
- Allais, M. (1953). Le comportement de l’homme rationnel devant le risque: critique des postulats et axiomes de l’Ecole Americaine. *Econometrica*, 21(4), 503–546. <https://doi.org/10.2307/1907921>.
- Arthur, W.B., Holland, J.H., LeBaron, B., Palmer, R., & Tayler, P. (1997). Asset pricing under endogenous expectations in an artificial stock market. In Arthur, W.B., Durlauf, S.N., & Lane, D.A. (1st ed.), *The Economy as an Evolving Complex System II*, 15-44. CRC Press. <https://doi.org/10.1201/9780429496639>.
- Benhammada, S., Amblard, F., & Chikhi, S. (2017). An artificial stock market with interactions network and mimetic agents. *9th International Conference on Agents and Artificial Intelligence*. 390-397. <https://doi.org/10.5220/0006118803900397>.
- Benink, A.H., Gordillo, J.L., Pardo, J.P., & Stephens, C.R. (2010). Market efficiency and learning in an artificial stock market: A perspective from Neo-Austrian economics. *Journal of Empirical Finance*, 17(4), 668-688. <https://doi.org/10.1016/j.jempfin.2010.02.004>.
- Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22(1), 23–36. <https://doi.org/10.2307/1909829>
- Booth, A. D. (1967). A “law” of occurrences for words of low frequency. *Information and control*, 10(4), 386-393. [https://doi.org/10.1016/S0019-9958\(67\)90201-X](https://doi.org/10.1016/S0019-9958(67)90201-X)
- Bradford, S.C. (1934). Sources of information on scientific subjects. *Engineering*, 137, 85-86.

- Chen, H., & Yeh, C. (2002). On the emergent properties of artificial stock markets: the efficient market hypothesis and the rational expectations hypothesis. *Journal of Economic Behavior & Organization*, 49, 217-239. [https://doi.org/10.1016/S0167-2681\(02\)00068-9](https://doi.org/10.1016/S0167-2681(02)00068-9).
- Chen, S., & Huang, Y. (2008). Relative risk aversion and wealth dynamics. *Information Sciences*, 177(5), 1222-1229. <https://doi.org/10.1016/j.ins.2006.08.007>.
- Chen, S., & Huang, Y. (2008). Risk preference, forecasting accuracy and survival dynamics: Simulations based on a multi-asset agent-based artificial stock market. *Journal of Economic Behavior & Organization*, 679 (3-4), 702-717. <https://doi.org/10.1016/j.jebo.2006.11.006>.
- Chung, K.H., & Cox, R.A.K. (1990). Patterns of productivity in the finance literature: a study of the bibliometric distributions. *Journal of Finance*, 45(1), 301-309. <https://doi.org/10.1111/j.1540-6261.1990.tb05095.x>
- Chen, X., & Tokinaga, S. (2004). Multi-agent-based modeling of artificial stock markets by using the co-evolutionary GP approach. *Journal of the Operations Research Society of Japan*, 47(3), 163-181. <https://doi.org/10.15807/jorsj.47.163>
- Ducha, F.A., Atman, A.P.F., & Magalhães, A.R.B. (2021). Information flux in complex networks: Path to stylized facts. *Physica A: Statistical Mechanics and its Applications*, 566, 1-10. <https://doi.org/10.1016/j.physa.2020.125638>.
- Fama, E. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, 25, 383-417. <http://dx.doi.org/10.2307/2325486>.
- Fama, E. (1991). Efficient capital markets: II. *The Journal of Finance*, 46(5), 1575-1618. <https://doi.org/10.1111/j.1540-6261.1991.tb04636.x>
- Goffman, W. (1971). A mathematical method for analyzing the growth of a scientific discipline. *Journal of the Association for Computing Machinery*, 18(2), 173-185. <https://doi.org/10.1145/321637.321640>
- He, X., Li, Y., & Zheng, M. (2019). Heterogeneous agent models in financial markets: A nonlinear dynamics approach. *International Review of Financial Analysis*, 62, 135-149. <https://doi.org/10.1016/j.irfa.2018.11.016>
- Holland, J.H. (1992). Complex adaptive systems. *Daedalus*, 121(1), 17-30. <https://www.jstor.org/stable/20025416>
- Kahneman, D., & Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2), 263-291. <https://doi.org/10.2307/1914185>
- Kluger, B.D., & McBride, M.E. (2011). Intraday trading patterns in an intelligent autonomous agent-based stock market. *Journal of Economic Behavior & Organization*, 79(3), 226-245. <https://doi.org/10.1016/j.jebo.2011.01.032>.
- Kwapien, J., & Drozd, S. (2012). Physical approach to complex systems. *Physics Reports*, 515, 115-226. <https://doi.org/10.1016/j.physrep.2012.01.007>.
- LeBaron, B., Arthur, W.B., & Palmer, R. (1999). Time series properties of an artificial stock Market. *Journal of Economic Dynamics and Control*, 23, 1487-1516. [https://doi.org/10.1016/S0165-1889\(98\)00081-5](https://doi.org/10.1016/S0165-1889(98)00081-5)
- Lotka, A.J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12), 317-323. <https://www.jstor.org/stable/24529203>
- Mandes, A. (2020). Impact of electronic liquidity providers within a high-frequency agent-based modeling framework. *Computational Economics*, 55, 407-450. <https://doi.org/10.1007/s10614-019-09891-1>
- Mandes, A., & Winker, P. (2017). Complexity and model comparison in agent based modeling of financial markets. *Journal of Economic Interaction and Coordination*, 12, 469-506. <https://doi.org/10.1007/s11403-016-0173-0>

- Manahov, V. (2016). A note on the relationship between high-frequency trading and latency arbitrage. *International Review of Financial Analysis*, 47, 281-296. <https://doi.org/10.1016/j.irfa.2016.06.014>.
- Manahov, V. (2016). Can high-frequency trading strategies constantly beat the market? *International Journal of Finance and Economics*, 21, 167-191. <https://doi.org/10.1002/ijfe.1541>.
- Manahov, V. (2016). Front-running scalping strategies and market manipulation: Why does high-frequency trading need stricter regulation? *Financial Review*, 51, 363-402. <https://doi.org/10.1111/fire.12103>
- Manahov, V., Hudson, R., & Hoque, H. (2015). Return predictability and the “wisdom of crowds”: Genetic programming trading algorithms, the marginal trader hypothesis and the Hayek hypothesis. *Journal of International Financial Markets*, 37, 85-98. <https://doi.org/10.1016/j.intfin.2015.02.009>.
- Oldham, M. (2020). Quantifying the concerns of Dimon and Buffett with data and computation. *Journal of Economic Dynamics and Control*, 113, 65-89. <https://doi.org/10.1016/j.jedc.2020.103864>.
- Pao, M. L. (1978). Automatic text analysis based on transition phenomena of word occurrences. *Journal of the American Society for Information Science*, 29(3), 121-124. <https://doi.org/10.1002/asi.4630290303>
- Patel, R., Goodell, J. W., Oriani, M.E., Paltrinieri, A., & Yarovaya, L. (2022). A bibliometric review of financial market integration literature. *International Review of Financial Analysis*, 80, 1-18. <https://doi.org/10.1016/j.irfa.2022.102035>.
- Ponta, L., Pastore, S., & Cincotti, S. (2018). Static and dynamic factors in an information-based multi-asset artificial stock market. *Physica A: Statistical Mechanics and its Applications*, 492, 814-823. <https://doi.org/10.1016/j.physa.2017.11.012>.
- Raberto, M., Cincotti, S., Focardi, S.M., & Marchesi, M. (2003). Traders' long-run wealth in an artificial financial market. *Computational Economics*, 22, 255-272. <https://doi.org/10.1023/A:1026146100090>.
- Rekik, Y.M., Hachicha, W., & Boujelbene, Y. (2014). Agent-based modeling and investors' behavior explanation of asset price dynamics on artificial financial markets. *Procedia Economics and Finance*, 13, 30-46. [https://doi.org/10.1016/S2212-5671\(14\)00428-6](https://doi.org/10.1016/S2212-5671(14)00428-6).
- Ryu, D., Yang, H., & Yu, J. (2021). Insider trading and information asymmetry: Evidence from the Korea Exchange. *Emerging Markets Review*, In Press <https://doi.org/10.1016/j.ememar.2021.100847>.
- Staccioli, J., & Napoletano, M. (2021). An agent-based model of intra-day financial markets dynamics. *Journal of Economic Behavior & Organization*, 182, 331-348. <https://doi.org/10.1016/j.jebo.2020.05.018>.
- Schmitt, N., & Westerhoff, F. (2017). Heterogeneity, spontaneous coordination and extreme events within large-scale and small-scale agent-based financial market models. *Journal of Evolutionary Economics*, 27, 1041-1070. <https://doi.org/10.1007/s00191-017-0504-x>
- Tesfatsion, L. (2016). *Review of agent-based computational economics: how the idea originated and where it is going*. Iowa State University, Department of Economics. <https://ideas.repec.org/p/isu/genstf/201612010800001695.html>
- Tsao, C.Y., & Huang, Y.C. (2018). Revisiting the issue of survivability and market efficiency with the Santa Fe Artificial Stock Market. *Journal of Economic Interaction and Coordination*, 13, 537-560. <https://doi.org/10.1007/s11403-017-0192-5>

- Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323. <https://doi.org/10.1007/BF00122574>.
- Vanfossan, S., Dagli, C.H., & Kwasa, B. (2020). An agent-based approach to artificial stock market modeling. *Procedia Computer Science*, 168, 161-169. <https://doi.org/10.1016/j.procs.2020.02.280>.
- Wamba, S.F., Bawack, R.E., Guthrie, C., Queiroz, M.M., & Carillo, K.D.A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technological Forecasting and Social Change*, 164, 1-27. <https://doi.org/10.1016/j.techfore.2020.120482>.
- Wang, J., Wu, C., & Zhong, X. (2021). Prospect theory and stock returns: Evidence from foreign share markets. *Pacific-Basin Finance Journal*, 69, 1-24. <https://doi.org/10.1016/j.pacfin.2021.101644>.
- Westerhoff, F. (2006). Technical analysis based on price-volume signals and the power of trading breaks. *International Journal of Theoretical and Applied Finance*, 9(2), 227-244. <https://doi.org/10.1142/S0219024906003512>
- Yang, X., Zhang, J., & Ye, Q. (2020). Tick size and market quality: Simulations based on agent-based artificial stock markets. *Intelligent Systems in Accounting, Finance and Management*, 27, 125 – 141. <https://doi.org/10.1002/isaf.1474>.
- Yeh, C.H. (2007). The role of intelligence in time series properties. *Computational Economics*, 30, 95–123. <https://doi.org/10.1007/s10614-007-9089-z>
- Yeh, C.H., & Yang, C.Y. (2010). Examining the effectiveness of price limits in an artificial stock market. *Journal of Economic Dynamics and Control*, 34(10), 2089-2108. <https://doi.org/10.1016/j.jedc.2010.05.015>.
- Yeh, C. H., & Yang, C.Y. (2013). Do price limits hurt the market? *Journal of Economic Interaction and Coordination*, 8, 125–153. <https://doi.org/10.1007/s11403-012-0107-4>.
- Zhang, J., McBurney, P. & Musial, K. (2018). Convergence of trading strategies in continuous double auction markets with boundedly-rational networked traders. *Review of Quantitative Finance and Accounting*, 50, 301–352. <https://doi.org/10.1007/s11156-017-0631-3>
- Zipf, G.K. (1949). *Human behavior and the principle of least effort: an introduction to human ecology*. Addison-Wesley Press. [https://doi.org/10.1002/1097-4679\(195007\)6:3<306::AID-JCLP2270060331>3.0.CO;2-7](https://doi.org/10.1002/1097-4679(195007)6:3<306::AID-JCLP2270060331>3.0.CO;2-7)