Applications of genetic programming to finance and economics: past, present, future

Anthony Brabazon[1] · Michael Kampouridis[2] · Michael O'Neill[1]

Abstract

While the origins of genetic programming (GP) stretch back over 50 years, the field of GP was invigorated by John Koza's popularisation of the methodology in the 1990s. A particular feature of the GP literature since then has been a strong interest in the application of GP to real-world problem domains. One application domain which has attracted significant attention is that of finance and economics, with several hundred papers from this subfield being listed in the GP bibliography. In this article we outline why finance and economics has been a popular application area for GP and briefly indicate the wide span of this work. However, despite this research effort there is relatively scant evidence of the usage of GP by the mainstream finance community in academia or industry. We speculate why this may be the case, describe what is needed to make this research more relevant from a finance perspective, and suggest some future directions for the application of GP in finance and economics.

1 Introduction

A particular feature of the field of genetic programming (GP) is that there has always been strong research interest in applying GP methodologies to real-world problems. One of GP's main advantages is its ability to capture domain-specific representations, allowing for a much more efficient search. Perhaps this is unsurprising as the concept of 'automatic programming' or as put more colloquially by John Koza 'getting computers to solve problems without explicitly programming them' is inherently concerned with problem-solving. Indeed, of the five papers published

✉ Michael Kampouridis
M.Kampouridis@kent.ac.uk

1   University College Dublin, Dublin, Ireland

2   University of Kent, Medway, UK

in the very first issue of GPEM, three could broadly be considered as being applications-focussed with [63] commenting in 2010 that '...[in] recent years the field has become increasingly applied...'.

The GP bibliography contains over 13,400 references as of October 2018 and several hundred of these papers concern applications in finance and economics. This strand of literature has attracted considerable interest with a review by [56] for the 10th anniversary edition of GPEM noting that more than half of the top 20 most-downloaded papers from the GP bibliography (http://www.cs.bham.ac.uk/~wbl/biblio/) in the period September 2006–October 2009 concerned finance. Financial application papers have also been popular at leading EC/GP conferences, frequently featuring at GECCO, EuroGP, the annual track on Evolutionary Computation in Finance and Economics at the IEEE Congress on Evolutionary Computation, the IEEE Conference on Computational Intelligence for Financial Engineering (CIFEr), and EvoBAFIN held annually as part of EvoApps.

The interest in the application of GP to problems in finance and economics has mirrored a wider trend in the field of computational intelligence in the late 1980s and early 1990s when a plethora of studies emerged applying (initially) neural networks for financial forecasting purposes and subsequently evolutionary approaches. In the latter case, attention was at first focused on the application of GAs for model parameter optimisation and variable selection [9, 35]. GAs quickly started being applied to other domains, too, such as portfolio optimisation. As Evolutionary Automatic Programming (EAP) methodologies such as GP were popularised [53, 55] they too were applied for economic modelling, financial forecasting [62], and trading system induction [7]. It is an interesting footnote in the history of GP that one of the earliest exemplar applications of the methodology by John Koza was to recover the well-known *exchange equation M = PQ/V* which relates the money supply (*M*), price level (*P*), gross national product (*Q*), and velocity of money (*V*) in an economy [53, 54].

In this paper we seek to address a key question. In spite of the significant research effort dating back to the early days of GP concerning the application of GP to the domain of economics and finance at EC/GP conferences and in EC/GP journals, there has been surprisingly little work published using GP in the mainstream finance academic community or indeed publicly disclosed by industry practitioners.[1] The question arises as to why this is so? In the earliest days it could have been suggested that knowledge of GP had not yet spread to the finance community. However, given that it is now nearly 20 years since the first paper using GP appeared in the mainstream finance literature [7], this argument does not seem particularly plausible.

The remainder of this contribution is organised as follows. In Sect. 2 we briefly contextualise a sample of the extant literature in finance which adopts a GP approach. Section 3 outlines some of the shortcomings of this work, leading to

---

[1] Of course, industry participants have a good reason to keep successful applications of new technologies secret and this could explain the relative lack of industry practitioners that discuss the use of GP and other advanced methodologies. There are a few notable exceptions, such as Sentient Technologies, which has used evolutionary and deep learning for areas such as e-commerce and trading.

Sect. 4 which provides some specific suggestions as to what needs to be done to ensure that future research has credibility to domain experts. We also provide some suggestions for future areas of research. Finally, Sect. 5 concludes this article.

2 Archaeology of finance applications using GP

Much of the early GP research in finance was primarily proof of concept in nature, being constrained by the availability of computing power and data. Nonetheless, this work laid the foundation for more comprehensive subsequent studies which also took advantage of the increasing maturity of GP methodologies. In this section, we initially review the reasons as to why GP can be a valuable methodology in the domain of economics and finance, then we illustrate the span of research undertaken in this domain.

2.1 Why apply GP in finance and economics?

A key feature of GP is its ability to co-evolve both the solution form and the relevant solution parameters. As a result, GP can simultaneously create new programs (solutions) and also optimise their parameters. This offers particular benefit in finance and economics as we often lack strong theoretical models which well-explain phenomena of interest. Although there may be a multitude of potential explanatory variables, the relationships amongst these variables which may be non-linear [13], are often poorly-understood. Therefore, a powerful model induction methodology, such as GP, can be particularly advantageous.

Another strong feature of GP is that it permits the incorporation of domain knowledge where this exists, leading to the generation of 'solutions' of a particular or desired form. The application of GP operators, such as crossover and mutation, on such 'solutions' that have incorporated domain knowledge can then make the search more efficient. Grammar-based forms of GP are particularly suited to the easy incorporation of domain knowledge and/or the placing of a clear structure on the to-be evolved solutions.

An additional oft-claimed benefit of GP in financial applications is that, unlike black box approaches to model induction such as neural nets, GP holds out the potential of generating human-readable solutions. In turn, such solutions could potentially generate new insights and lead to theory development. Human-readability can be important for financial applications as in some cases there may be a regulatory requirement to be able to explain decisions (such as the decision whether or not to grant a loan). In addition, from a risk management perspective, the appetite of financial institutions (or indeed regulators) for black-box systems which make automated decisions can be limited.

Furthermore, evolutionary algorithms, such as GP, are forgiving in terms of calculation of the fitness function. Due to the manner in which selection can operate, we do not need to have absolute or precise measure of solution quality for the algorithm to effectively search and optimise the space of financial or economic models

under investigation. Rather, coarse-grained calculations or even approximations of fitness can work, such that the algorithm only requires relative performance of each individual to other population members. This can be attractive in environments which are often noisy or where we have incomplete theory, and an attempt to adopt a precise simulation could be very computationally expensive. Of course relative quality goes hand-in-glove with the idea of being good enough. In the Business world it is said on occasion "the perfect is the enemy of the good". In other words, survival and prosperity is not necessarily dependent upon finding the global optimal solution; finding and adopting solutions which are better than our competitors is good enough. This is consistent with adopting global search and optimisation heuristics such as GP.

Also relating to fitness, there are well-known strategies in evolutionary computation to handle multiple/many competing objectives, which often exist in real world application domains, such as financial modelling; for example, when developing a trading strategy it is critical to balance risk and reward, where often greater risk can lead to larger rewards, but at the risk of significant loses.

Finally, due to advances in computer hardware and the availability of high-quality electronic financial information, the practicality of applying GP in finance and economics has increased.

2.2 Taxonomy of existing literature

In the following section we illustrate a sample of the literature at the nexus of GP and finance in order to highlight the diversity of this research. This will also motivate the discussion in Sect. 3, where we will discuss shortcomings of much of this work. We do not seek to provide a comprehensive review of this voluminous literature and readers interested in this are referred to [16, 21, 71] for more details on individual studies and a description of how each area of application has developed over time. Figure 1 illustrates some of the main application areas.

2.2.1 Forecasting

Financial forecasting can involve the prediction of future values of macroeconomic variables, individual stock value, market indices, commodity futures, or the volatility of some financial instrument, amongst many other possibilities. It is a heavily researched topic in mainstream finance and unsurprisingly, it is also one of the most popular areas in terms of GP application to economics and finance. Many different problems have been explored, such as forecasting different economic and financial variables, e.g., GDP and the price of a stock [44, 46, 72]. Perhaps the most popular area in financial forecasting (in terms of volume of publications) is evaluating the profitability of technical trading rules (also known as trading strategies). Many such studies have appeared over the years, focusing on different markets (e.g., foreign-exchange markets [61], stock markets [7, 51]), using many different types of data (daily [47, 50], monthly [40], and more recently intra-day [36] and event-based
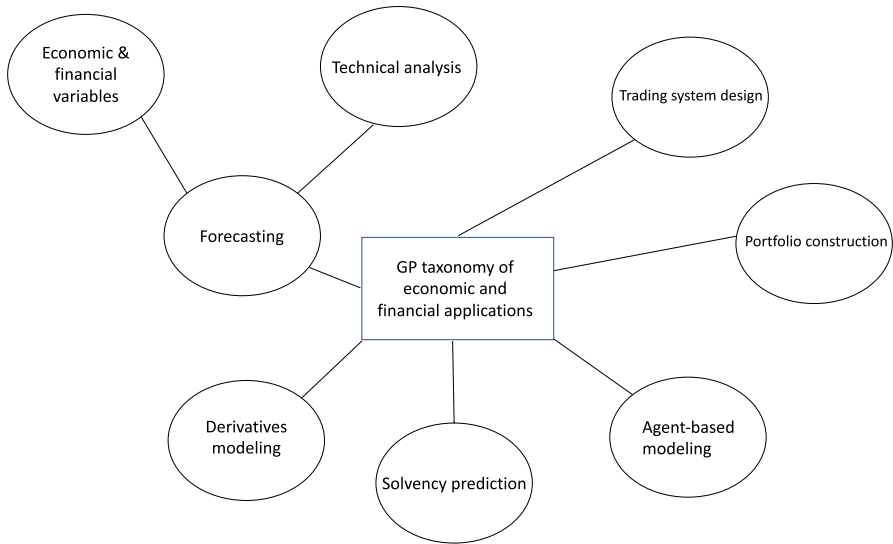
Fig. 1 GP taxonomy of economic and financial applications

data [1]). There have also been studies combining technical analysis with fundamental and macroeconomic analysis [25].

2.2.2 Trading system design

A significant number of applications of GP in finance have concerned trading system design. Typically, systems take one of two approaches, using either fundamental data (e.g., data from accounting statements, industry-level data, or macroeconomic data) or market-based data from an electronic order book such as price/volume data. More recently, some studies have sought to integrate data from multiple time horizons [25] in making investment decisions. A significant issue in the design of these systems is the appropriate choice of objective/fitness function, as poor choices will result in systems which do not perform as required, undermining the credibility of the utility of GP as a method.[2]

A related, important, issue in trading financial assets is the efficient execution of large institutional size orders, and applications of GP for this task have begun to emerge in recent years. When trading shares, particularly when an investor is looking to buy or sell a large quantity of stock, the problem of *market impact* arises. Market impact occurs when the actions of an investor start to move the price adversely against themselves. Hence, market impact is the difference between a transaction price and what the market price would have been in the absence of the transaction. The design of a trade execution strategy therefore seeks to trade off the

---

[2] A more detailed discussion on the importance of the appropriate selection of fitness functions takes place in Sect. 3.5.

cost of trading quickly (i.e., market impact) against the cost of trading slowly (risk that market price moves over time and/or risk that trade never gets completed).

A novel approach was taken by [32, 33] where grammatical evolution (GE) was used to evolve a dynamic trade execution strategy, with the resulting rule adapting to changing market conditions. Based on the finance literature analysing the relationship between order placement and the information content of limit order books, six order book metrics were selected as potential inputs for an execution strategy. In the implementation, GE was found to be able to evolve quality trade execution strategies and its results proved highly competitive against two basic benchmark execution strategies. A detailed discussion of the application and the relevant background finance literature is provided in [34].

2.2.3 Portfolio construction

In finance a portfolio refers to a grouping of financial assets such as stocks, bonds and cash equivalents, as well as their mutual, exchange-traded and closed-fund counterparts. Portfolio management involves the art and science of making decisions about investment mix and policy, matching investments to objectives, asset allocation and balancing risk and return. GP has proven popular for stock selection for investment portfolios. Early work in this field includes [11], which used GP to evolve filter rules to select stocks for the high-technology manufacturing industry. Some work has also focused on promoting generalisation and avoiding overfitting [12], and on selecting stocks in complex environments (e.g., environments that consist of a range of different market dynamics, such as bull, bear, and sideways market conditions) [75].

2.2.4 Derivatives modelling

Derivatives are contracts whose value is derived from the value of underlying assets, such as equities, interest rates, currencies, market indices, commodities etc. Two of the best known forms of derivative are futures and options. A key issue for investors wishing to trade in derivatives is the determination of the fair price for the financial instrument. For some standard derivatives (based on specific assumptions such as continuous time finance theory), closed-form pricing equations have been determined (e.g., the Black–Scholes model [15, 60] for pricing European options, the Cox et al. binominal model [26] etc.). In reality, some of the key assumptions of these models do not hold in real-world option markets, and hence they do not explain observed option prices correctly. Due to the complexity in developing closed form theoretical models for options pricing, the domain is particularly amenable to model-induction techniques. GP allows both the structure and the parameters of the pricing model to be estimated directly from the data, thereby extracting the pricing model implicitly.

GP has been used for developing option pricing models in the context of derivatives price modeling [23, 24, 76], and the development of hedging strategies which make use of derivative instruments [77, 78]. More recently, GP has been

used for the relatively new area of derivatives, namely weather derivatives, which are financial instruments for dealing with adverse weather conditions [3–5, 27, 29, 30].

2.2.5 Solvency modelling

Solvency modelling is an important practical problem and has a long history in mainstream finance. One of the earliest strands of this concerned corporate bankruptcy prediction with later work looking at the closely related problem of prediction of corporate and bond ratings from agencies such as Moody's. GP has been applied to all of these areas with, for example, [58] using GP to classify Norwegian firms into soon-to-be-bankrupt and non-bankrupt classes, and [65] predicting insurance companies' insolvency. GP has also been applied for credit-scoring of potential customers [64].

2.2.6 Agent-based modelling

Agent-based modelling (ABM) allows the simulation of markets which consist of heterogeneous agents, with differing risk attitudes and differing expectations to future outcomes, in contrast to traditional assumptions of investor homogeneity and rational expectations. ABM attempts to explain market behaviour, replicate documented features of real-world markets, and allows us to gain insight into the likely outcomes of different regulatory policy choices. The essence of ABM lies in the notion of autonomous agents whose behaviour evolves endogenously leading to complex, emergent, system dynamics which are not predictable from the properties of the individual agents.

In developing an agent-based model, a key question is how do agents learn and adapt their strategies over time. These 'learning mechanisms' can be implemented in many ways including use of a neural network or by means of an evolutionary approach where each agent's behaviour or strategy adapts over time in response to environmental feedback.

A significant literature using GP in ABM has emerged. For example, [22] created an agent-based artificial stock market and found evidence supporting the Efficient Market Hypothesis, [59] modeled different types of traders (e.g., fundamental, technical, and noise traders) in an artificial stock market, and [48] modeled elements of market microstructure. Other applications of ABM include the simulation of a foreign exchange market [45], the modelling of an artificial stock option market [37] and the modelling of an artificial payment card market [6]. As it is not within the scope of this paper to provide a thorough review on the topic, we refer the reader to [19, 20], which provide in-depth reviews.

As can be seen from the brief illustration above, the span of applications of GP in finance and economics is very wide. In the next section we describe some of the shortcomings exhibited by much of this work.

3 Shortcomings of prior work

Amongst the articles in the 10th anniversary issue of GPEM a paper by [63] focussed on 'open issues' in GP. The authors noted that '...GP does not seem to be universally recognised as a mainstream and trusted problem solving strategy, despite the fact that in the last decade GP has been shown to outperform some of these methods [i.e., other machine learning methods] ...'.

A decade on, despite the considerable research literature applying GP to problems in finance and economics, a similar comment can still be made. GP has not yet found widespread application in the finance industry, or indeed in the mainstream finance research literature. For example, a lengthy 280 page report [52] produced by J. P. Morgan aimed to provide a framework for Machine Learning and Big Data investing. This report also included an overview of several Machine Learning methods. Despite containing detailed coverage of algorithms such as k-nearest neighbours, support vector machines, random forest, and neural networks, it is notable that GP was completely omitted. Similarly, Coursera,[3] and Udacity,[4] explicitly focus on algorithms such as deep learning and support vector machines, as part of their online course portfolio for finance (Coursera: Introduction to Fundamentals of Machine Learning in Finance, Udacity: Artificial Intelligence for Trading). Once again, GP is not mentioned.

This raises the obvious question as to why this is the case? Ten years ago it could be argued that this lack of real-world application was partly due to the relative novelty of GP methods and need to allow time for knowledge of GP to disseminate into other domains including finance. A decade on, this argument holds even less water and it is now fully 20 years since the first paper using GP appeared in the mainstream finance literature [7].

For the rest of this section, we will identify the reasons we believe that GP has not been more widely applied. In large measure this stems from the lack of rigor of much of this research. To frame a discussion of this, inspired by [63], we also present some open issues that have had implications to finance and economics applications, explaining how these issues have contributed to shortcomings of prior work.

3.1 Disconnect between computer science and finance

Generally, mainstream research in finance is strongly framed by theory. Building on prior work, a theoretical framework is proposed which identifies plausible explanatory variables and a model structure. Hypotheses are developed based on this, and finally these are tested using real-world data. It is usually required that a study makes a clear contribution to theory. In contrast, GP's appeal (and machine learning in general) is discovering its own models from raw data; GP is thus most usually applied

---

in an inductive manner, where a series of potential explanatory variable (terminals) are proposed and GP is used to evolve a model which links these.

A clear disconnect has opened up between the computer science and finance literatures with many papers applying GP to finance appearing in the former, but very few appearing in the latter. Many of the papers appearing in the computer science literature exhibit little real linkage with the finance literature, tending to focus on the 'method' rather than clearly addressing core research questions in the finance domain. Hence, the finance data is used as a test bed but there is little real contribution to finance theory or practice. Results from these studies are not credible from a finance perspective. Notably, a significant number of these papers have few or no co-authors with finance expertise.

A typical example of this situation can be found in a number of papers from the late 1990s, which were applying GP for trading in the stock market; while these studies would conclude that they were obtaining profitable results, they would also admit that their strategies were not able to beat the buy-and-hold strategy. In reality, the duration of the datasets, starting in the mid-90s, and the dot-com bubble of the time, made it hard for any algorithm to beat buy and hold.

3.2 Problem scalability

Many published studies using GP in finance are 'proof of concept' and the research is never progressed to (hard) real-world exemplars. Often such studies make unrealistic assumptions, e.g., omitting transaction costs when an action takes place in the market. Hence, while GP can perform well on these toy problems, it is questionable or simply unknown how well the method will scale to problems of increasing complexity and difficulty, or when fewer (or no) assumptions are made. This hurts the credibility of GP's ability to tackle difficult financial and economics problems.

3.3 Problem benchmarks

A major issue in the literature concerning GP and finance/economics is a plethora of papers claiming to 'outperform' straw man benchmarks, such as buy-and-hold in a trading framework. Studies are typically idiosyncratic adopting datasets from differing markets and time periods, using differing fitness functions, and differing GP system parameters. This makes it impossible to meaningfully compare results between studies. There is no tradition of use of agreed problem benchmarks in financial and economic applications.

This is by no means a problem only for the GP community. However, if we truly want our algorithms' results to be understood and appreciated by the finance and economics community, we need to start building these benchmarks. It is not an easy task, as there are many different applications, as we discussed in Sect. 2. Nevertheless, we should focus towards agreeing on certain benchmarks for each application. Competitions in conferences, as it already happens in GECCO and CEC, could become mainstream, and studies could aim at verifying their performance in such competitions, to increase the credibility of their results.

3.4 Transparency of GP models

Although the use of GP holds out the possibility of generating human-readable models, few basic implementations of GP will actually produce models which are amenable to easy interpretation. In all but the simplest applications, the resulting models (GP individuals) tend to be complex and this can make it difficult to gain any significant insight into the potential underlying model and therefore to contribute to theory development. Another practical issue that arises is that black (or grey) box models are often impractical to implement in real-world finance settings, as for example there can be regulatory requirements to 'explain' the rationale for decisions such as why an individual was denied credit. Similarly, for risk management purposes, risk committees in financial institutions can be slow to approve the use of systems whose decisions cannot be easily explained.[5]

One obvious approach to assisting GP in producing models which are more likely to be interpretable is to place some basic structure on the models being generated a priori. The use of grammar-based GP approaches can assist, with the modeller using domain knowledge to bound in the model space to be searched via the definition of a suitable grammar.

3.5 Fitness function design

One of the most critical decisions in applying GP, or indeed any evolutionary approach, to problems in finance and economics is the appropriate choice of fitness function. Poor choices will result in systems which do not perform as required, undermining the credibility of the utility of GP as a method. For example, selection of raw profit as a fitness metric in a trading algorithm is quite likely to lead to the generation of trading systems with undesirable risk characteristics, in terms of the variance or skewness of trading returns; this is because a fitness metric which focusses on raw returns contains no penalty against trading systems which produce volatile returns [17, 18].

Another problem with traditional approaches to fitness function selection is that risk metrics, such as variance, do not take account of the temporal ordering of returns. A sequence of negative returns can lead to large drawdowns which can have detrimental consequences if an individual investor or fund runs out of capital, suggesting an important role for consideration of all aspects of the shape of the produced equity curve in assessing trading systems.

A key, and often under utilised, strength of GP relative to traditional modelling techniques such as regression is that the fitness function can easily be tailored for the problem of interest. For example, in trading context, GP permits the implementation of a fitness metric which will produce any desired returns distribution—rather than

---

[5] There are some exceptions, e.g. high-frequency trading hedgefunds, where black box models are becoming more acceptable, especially due to the good performance of algorithms such as deep learning. Nevertheless, the problem remains that there are many other areas in economics and finance that black (or grey) box models are impractical to implement.

just concentrating on the first and second moments (mean and variance of returns respectively) of this distribution. Unfortunately, many GP applications in finance and economics fail to make appropriate use of this flexibility and implement simple, unrealistic, fitness functions which undermine the credibility of the study's results.

Another credibility-related issue is the problem of 'data snooping'. When a dataset is used multiple times for model selection, an apparently good model could occur due to chance alone rather than representing a truly robust model of the data-generating process [74]. In these cases, the model will likely perform poorly out of sample. Data snooping is a particular concern in powerful methodologies such as GP, due to the very large number of models that can be generated and tested against the same dataset during training. Although there have been studies addressing this issue (e.g., [2, 67]), many studies applying GP in finance and economics do not adequately take data snooping issues into account.

3.6 GP usability

As already noted, knowledge of GP amongst finance practitioners, or indeed, amongst finance academics is quite low. Unlike traditional modelling techniques such as regression, GP is more difficult to apply in a 'canned fashion'. Obtaining good results from a GP system usually requires the user to carefully define parameters and attributes including the function set, terminal set, fitness function, appropriate diversity generation operators, appropriate problem representation, and so on. It is also important to carefully protect against the tendency of low grade implementations of GP to overfit training data and therefore generalise less well out of sample. These issues make it more difficult to use GP and underscores the need for multi-disciplinary teams in order to ensure a quality result when applying GP in the finance/economics domain.

In summary, while GP has significant attributes which lend it well to finance and economics applications, the realisation of its full potential requires a new mindset, that of application of the 'right methodology' to the 'right problem' within a multi-disciplinary setting. Unfortunately, much of the published work at the nexus of GP and finance/economics is too method focussed, and has little serious credibility in the domains of finance and economics. A more rigorous, multi-disciplinary, approach to the peer-review of these application papers is required in order to ensure that GP gains wider acceptance as a truly-valued methodology in finance and economics.

4 Looking to the future

Despite the shortcomings of much prior work applying GP in a financial setting, automatic programming methodologies hold continuing promise for significant research in finance and economics in the future. Trends which support the use of GP in financial applications include:

- the increasing availability of data (including non-traditional, finance-related, data) in electronic form,
- the increasing power of computers which alters the relative costs of financial theory versus inductive modeling methodologies,
- courses in computational finance and economics are becoming staple offerings in leading colleges, and
- the Red Queen effect

In this section we initially indicate a number of interesting research opportunities to which GP could be applied in coming years. First of all, we discuss the fact that in addition to having more and more data available, recently there have been opportunities to explore new types of data, such as 'digital footprints', text data, and social media data. This takes place in Sect. 4.1, where we consider these different types of data. Then, in Sect. 4.2, we shift our attention to agent-based modelling, as this is an area of continuous interest in finance and economics, and has wide applicability to many different problems in the field. We thus tease out some opportunities where GP could further advance the domain. Furthermore, Sect. 4.3 presents some financial application areas we believe will be in the frontiers of GP research in the near future. Finally, in Sect. 4.4 we discuss the future of GP research in economics and finance. From a review of current literature it is apparent that there is a requirement for a much more demanding 'review bar' to be set for future research work aimed at the nexus of GP and finance in order to boost the credibility of GP in the eyes of the finance community.

4.1 New data avenues

In this section we consider some of the opportunities for research arising from new data sources. In addition to the increasing availability of data, due to reasons related to improved hardware, the field is also seeing new data avenues being available, which can further improve our understanding of given problems.

Let us for instance consider a currently 'hot area' in finance, namely FinTech'—broadly defined here as any technological innovation in the financial sector. Many FinTech developments are focussing on consumer finance, typically looking at novel ways to deliver financial services cost effectively to mobile consumers. FinTech platforms are facilitating the disintermediation of financial services, and the entry of non-traditional institutions into the finance sector. Some interesting areas of current research concern the use of 'digital footprints' and 'customer behavioural profiling'—with each of these offering new ways to assess credit-worthiness amongst other potential finance applications. An example of the above can be found in a recent paper by [14], which used the 'digital footprint' from consumers in Germany to analyse the information content of an individual's footprint. Results showed that information people leave online when accessing a website has equal or greater predictive power as to a person's creditworthiness than their credit bureau score.

In addition, the last decade has seen an explosion in the variety of text data which is available for incorporation into financial models. While initial studies looked at

raw message count information, the next advance is to consider the content or 'sentiment' of these messages in order to assess whether investors are (un)favorably disposed towards a stock [38, 57]. Using text data in trading models has been made easier by the commercial availability of 'tagged' databases of financial news; for example, the Dow Jones Elementized News Feed places discrete pieces of news, keywords, timestamps, symbols and other crucial data, into XML-tagged fields for easy parsing and direct embedding into trading programs. Starting with the 'tagging' of text data in traditional financial newsfeeds, the range of available data has expanded to encompass social media, giving rise to the development of many systems which seek to embed 'sentiment' into financial prediction or trading models, in order to complement traditional data sources such as order book drawn from the market, financial statements or macroeconomic data.

Lastly, the mining of social media profiles has recently seen the development of individualised behavioral profiles for social media users which can potentially provide highly accurate information to (for example) lenders on a person's creditworthiness. These profiles can also be used to carefully target financial products to an appropriate demographic based on life stage, income and risk profile.

Only a relatively small number of papers have emerged applying GP in a finance setting to the above sources of data. This remains an exciting open area for future research.

4.2 Agent-based modelling

Agent-based modelling (ABM) is a tool which allows us to capture emergent phenomena and provide a natural description of a system, while doing this in a low cost and time-saving manner [10]. This need for dealing with emergent phenomena is never-ending (e.g., the case of Brexit being one of the most recent ones [66]), hence there will always be a need for further improving ABMs.

However, there are several challenges or pitfalls that ABMs face. Some well-known issues of ABMs are model calibration and how well the resulted model reflects the real world, e.g., are there salient properties missing in the ABM simulation or assumptions made that impact negatively on observed behaviours? In addition, simulation can be computationally expensive, and stochastic.

Furthermore, in designing ABMs, modellers face multiple design choices which can critically impact on the system's behaviour. If we take the example of designing a model of a financial market, important design questions which emerge include:

- Representation and structure of the actual trading agents. Agents can vary from simple budget constrained *zero intelligence agents* as in [39] to sophisticated learning agents as in [22].
- The actual mechanism that governs the trading of assets. Ways of designing this include assuming a simple price response to excess demand, building the market such that a kind of local equilibrium price can be found easily, or explicitly modelling the dynamics of trading to mimic the continuous trading of real-world markets.

- Types of securities to be incorporated into the agent-based market model, where typically simple securities (such as stocks) are considered.

Another important element of most ABMs is how agents process information and 'learn'. As noted in Sect. 2, GP has been applied to model agent learning, usually via the evolution of a behavioural 'rule'. The general approach has been to use 'vanilla' GP which can give rise to difficulties in the interpretation of the resulting 'behaviours' due to the complexity of the resulting evolved expressions. This approach also makes it difficult to incorporate a priori domain knowledge into agent behaviours.

A more powerful and natural approach would be to use grammar-based forms of GP such as grammatical evolution (GE). Although there have already been a number of studies applying grammar-based GP in an ABM framework ([32–34, 42, 43, 73] being a sampling of these), it is surprising that these approaches have not gained greater traction as a tool for ABM. Grammar-based GP facilitates modellers to place a desired structure (drawing on domain knowledge) via the grammar definition on the strategies that agents can employ, while still allowing considerable room for agents to adapt their strategies.

The ability to explicitly incorporate domain knowledge into agent strategies will also help ABM gain wider acceptance by the finance and economics community as it would allow a tighter integration of ABM designs with existing theory. In addition to modelling of agents in financial markets, there are a multiplicity of open opportunities for policy-focussed research via application of grammar-based GP approaches to ABM in economics.

4.3 Possible application areas

In this section we briefly discuss three popular application areas, which we believe will be in the frontiers of GP research in the near future based on their importance and potential impact in the financial world.

4.3.1 Hedging and derivative securities

Effective hedging of derivative securities is of paramount importance to derivatives investors. In reality, continuous rehedging is impossible on grounds of cost and market frictions. This raises the important practical question of when should a portfolio manager rebalance the portfolio? In practice, many portfolio managers employ relatively simple deterministic rebalancing strategies, such as rebalancing at uniform time intervals, or rehedging when the underlying asset moves by a fixed number of ticks. While such strategies are easy to implement, they will expose the portfolio to hedging risk, both in terms of timing and also as the strategies do not adequately consider market conditions.

While there has been some previous work on GP for hedging (e.g., [77, 78]), there is considerable scope for further research in this area. A useful area for future work could focus on using a GP-based hedging strategy with a joint objective function of maximising delta hedged returns whilst minimizing delta hedged risk.

Also, grammar-based GP could be highly beneficial for this application, as such approaches can incorporate domain knowledge from finance theory directly into the grammar.

4.3.2 Weather derivatives

As mentioned back in Sect. 2, a relatively new area of derivatives is weather derivatives. These are derivative products whose payoff is determined by the future values of weather metrics concerning phenomena such as temperature, snowfall, wind or rainfall.

Weather conditions affect the cash flows and profits of businesses in a multitude of ways. For example, energy companies (gas or electric) may sell fewer supplies if a winter is warmer than usual, leisure industry firms such as ski resorts, theme parks, hotels are affected by adverse weather outcomes on temperature, snowfall or rainfall, and agricultural firms can be impacted by weather conditions during the growing or harvesting seasons. The development of a market in weather derivatives allows firms to potentially hedge their exposure to adverse weather events.

A key input into a pricing model for weather risk (and associated derivatives) is a quality model of the distribution of the underlying weather metric at the location of interest. This is typically approximated using historical data for symbolic regression, but this task can be complicated by the existence of both short and long term weather variation and by local factors such as a heat-sink effect. Some initial studies have taken place, where GP has been used for the task of approximating the distribution of a weather metric at a specific location, and the consequent task of estimating a pricing model for a weather derivative (e.g., [3, 4, 28, 31]). Such studies have demonstrated the ability of GP to outperform the existing state-of-the-art financial or statistical techniques.

However, as weather derivatives are still in their infancy, there are a lot of opportunities for GP for further research. For example, in addition to temperature and rainfall that have currently been examined under GP frameworks, other types of weather derivatives could be examined, too, such as snow and wind. Given GP's strength for symbolic regression tasks, GP has the potential to make major impact in this field, first by making weather derivatives more popular and drawing new participants to the market, and second by introducing more accurate pricing models. In the long term, GP could also lead to a widely used pricing formulas, similarly to what happened after the introduction of the Black and Scholes formula [15].

4.3.3 Financial trading

Financial trading has and will continue to be an important area of application for GP. We can expect to see increasing sophistication in these applications with fuller implementation of smart entry and exit strategies, and greater attention being paid to market structure. There is a significant opportunity to undertake work concerning the design of appropriate fitness functions for trading applications. We can also expect to see greater integration of non-financial information, such as that from social media or official news wires into trading systems.

A particular area in financial trading that we believe will receive increased attention is event-based trading. Instead of observing price fluctuations from a physical time perspective (e.g. daily data, 10-min data, and so on), an alternative is to focus on significant events. Otherwise, there is a danger of missing important price movements, and thus potential profit opportunities. For instance, if we are using daily closing price summaries we would not be able to observe the 6 May 2010 Flash Crash, which was a United States trillion-dollar stock market crash that lasted for approximately 36 min.[6]

An event-based system is thus based on the idea that events can capture significant points in price movements that the traditional physical time methods cannot. A recent example of such an event-based system is Directional Changes (DC). Under this paradigm, the focus is on the size of price change, while time is the varying factor; on the other hand, under the physical-time paradigm time was fixed (e.g., daily closing prices). This new concept provides traders with new perspectives to price movements, and allows them to focus on those key points when an important event took place, blurring out other price details which could be considered irrelevant, or even noise.

While event-based systems are not completely new, the DC paradigm has been receiving increased attention lately in the area of financial forecasting (e.g., [8, 49, 69]) and this has led to the creation of new trading indicators which do not exist under physical-time price summaries [70]. This offers new opportunities for GP to take advantage of this largely unexplored research area. The only extant work using GP in this area is [1, 41], and an opportunity exists to extend this to create new trading strategies based on DC indicators, and to combine indicators from both physical and event-based time horizons. Furthermore, GP could be combined with other methods from machine learning and advanced analytics for advanced pattern recognition (e.g., deep learning); in this scenario, the patterns to detect would be the financial events, and GP could search the combinatorial space for a set of events, and even more sophisticated for sets of sequences of events. Being able to do this could potentially offer traders a much better understanding of the financial markets.

4.4 The future of GP research in economics and finance

After having presented some future research opportunities for GP, we would like to conclude this paper with a discussion of how we can, as a community, boost the credibility of GP in the eyes of the financial community. As discussed in Sect. 3, there are several shortcomings that affect this credibility. In the rest of this section, we re-visit these shortcomings and make suggestions on what needs to change.

In considering future directions for research at the nexus of computational intelligence and finance, the key development currently needing to take place is the maturing of research from simplistic 'proof of concept' studies to (hard) real-world finance problems. This transition requires the deepening of the realism of the

---

[6] http://blogs.wsj.com/marketbeat/2010/05/11/nasdaq-heres-our-timeline-of-the-flash-crash/ Last access: 10 October 2018.

financial problems addressed and the development of multi-disciplinary research teams with both GP and finance expertise.

This is crucial, because at the moment we see too many low quality papers, which do not offer anything new to the field. Due to a lack of rigour in reviewing processes and in spite of its promise, there is a real danger that GP will remain a marginal methodology in finance and economics. We echo the concerns of [68] who critiqued the proliferation of 'novel' metaheuristic methods, and the lack of rigorous interrogation to which 'new' methods are subjected.

Unfortunately, a significant number of papers applying GP in finance and economics are very marginal in their contribution, either applying an existing algorithm to a 'new' dataset, or making minor (unimportant) modifications to a previously published algorithm. It is also noticeable that many papers apply GP without providing a justification as to why it is the most appropriate algorithm for the task at hand. For example, if one wanted to connect a set of inputs to a set of outputs (as is the case in many of the existing publications), artificial neural networks might have been a more suitable algorithm than GP. Furthermore, although we argue that GP can produce white-box models, few published works make any serious attempt to interpret the final models produced. Lastly, a large number of published papers does not demonstrate that they are tackling a significant well-known financial or economics problem.[7]

To address the above issues, we believe that more rigorous standards should be applied when considering new studies for publication. At a minimum, we believe that submitted papers should:

- Establish that the paper is aiming to tackle a significant problem,
- Have a strong explanation why GP is suitable for the task,
- Demonstrate a significant contribution/novelty in the GP method employed,
- Make a serious attempt to interpret the final evolved models

Adopting a much more rigorous approach to the assessment of new research studies will in time enhance the credibility of GP to the finance community. Such standards already exist in the field, e.g. the "Humies" awards for human-competitive results produced by genetic and evolutionary computation[8] already enforce a set of standards. It is our belief that similar standards should be introduced when considering the publication of new GP studies applied to finance and economics.

Lastly, we would like to discuss the issue of benchmarks, which we view as a general problem of the domain, rather than only an issue of individual studies. As discussed in Sect. 3, the absence of benchmarks hinders GP's credibility. Although we understand this is not an easy issue to tackle, the community should nevertheless make efforts to address it. Competitions in conferences could play a vital role towards this goal. 'Enforcing' specific problems to be tested and using specific

---

[7] Of course, the significance of the problem should have been vetted by the scientific community; it shouldn't be left only to the authors of the paper to argue this.

[8] http://www.human-competitive.org/call-for-entries.

datasets would allow a 'common ground' for comparison. This would at least allow studies to increase the credibility of their results. In the long run, such competitions could also lead to agreed benchmarks for the different types of problems in finance and economics.

5 Conclusion

To sum up, this article discussed genetic programming as an application algorithm to the fields of finance and economics. We started by explaining why GP is an appropriate algorithm for this type of problems, and also illustrated the span of research undertaken in the domain. We then argued that GP has not found widespread application neither in the finance industry, nor in the mainstream finance research literature. We identified several issues which related to shortcomings to existing published works, and explained how these issues can be affecting GP's credibility to tackle difficult problems in finance and economics. Finally, in the last section of this work, we identified some areas that we believe GP holds continuing promise for impactful research in finance and economics in the years to come. Also, we presented some suggestions of how the research standards should change in the field in order to improve the quality of the research outputs.

References

1. A. Adegboye, M. Kampouridis, C.G. Johnson, Regression genetic programming for estimating trend end in foreign exchange market, in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)* (2017)
2. A. Agapitos, M. O'Neill, A. Brabazon, Evolutionary learning of technical trading rules without data-mining bias, in *Parallel Problem Solving from Nature, PPSN XI*, ed. by R. Schaefer, C. Cotta, J. Kołodziej, G. Rudolph (Springer, Berlin, 2010), pp. 294–303
3. A. Agapitos, M. O'Neill, A. Brabazon, Evolving seasonal forecasting models with genetic programming in the context of pricing weather-derivatives, in *Applications of Evolutionary Computation*, ed. by C. Di Chio (Springer, Berlin, 2012), pp. 135–144
4. A. Agapitos, M. O'Neill, A. Brabazon, *Genetic Programming for the Induction of Seasonal Forecasts: A Study on Weather Derivatives* (Springer, New York, 2012), pp. 159–188
5. A.K. Alexandridis, M. Kampouridis, S. Cramer, A comparison of wavelet networks and genetic programming in the context of temperature derivatives. Int. J. Forecast. **33**(1), 21–47 (2017)
6. B. Alexandrova-Kabadjova, E. Tsang, A. Krause, *Evolutionary Learning of the Optimal Pricing Strategy in an Artificial Payment Card Market* (Springer, Berlin, 2008), pp. 233–251
7. F. Allen, R. Karjalainen, Using genetic algorithms to find technical trading rules. J. Financ. Econ. **51**, 245–271 (1999)
8. A. Bakhach, E.P.K. Tsang, H. Jalalian, Forecasting directional changes in the fx markets, in *2016 IEEE Symposium Series on Computational Intelligence (SSCI)* (2016), pp. 1–8
9. R. Bauer, *Genetic Algorithms and Investment Strategies* (Wiley, New York, 1994)
10. A. Bazghandi, Techniques, advantages and problems of agent based modeling for traffic simulation. Int. J. Comput. Sci. Issues **9**(3), 115–119 (2012)
11. Y.L. Becker, P. Fei, A. Lester, *Stock Selection: An Innovative Application of Genetic Programming Methodology. Genetic Programming Theory and Practice IV* (Springer, Berlin, 2017)
12. Y.L. Becker, H. Fox, P. Fei, *An Empirical Study of Multi-objective Algorithms for Stock Ranking* (Springer, Boston, 2008), pp. 239–259

13. Y.L. Becker, U.M. O'Reilly, Genetic programming for quantitative stock selection, in *Proceedings of the First ACM/SIGEVO Summit on Genetic and Evolutionary Computation, GEC '09* (ACM, New York, 2009), pp. 9–16

14. T. Berg, V. Burg, A. Gombovic, M. Puri, On the rise of fintechs—credit scoring using digital footprints (July 10, 2018). Available at SSRN: https://ssrn.com/abstract=3163781 or http://dx.doi.org/10.2139/ssrn.3163781

15. F. Black, M. Scholes, The pricing of options and corporate liabilities. J. Polit. Econ. **81**(3), 637–654 (1973)

16. A. Brabazon, J. Dang, I. Dempsey, M. O'Neill, D. Edelman, *Natural Computing in Finance: A Review* (Springer, Berlin, 2012), pp. 1707–1735

17. A. Brabazon, M. O'Neill, *Biologically Inspired Algorithms for Financial Modelling* (Springer, Berlin, 2006)

18. R. Bradley, A. Brabazon, M. O'Neill, Objective function design in a grammatical evolutionary trading system, in *2010 IEEE World Congress on Computational Intelligence* (IEEE Press, Washington, DC, 2010), pp. 3487–3494

19. S.H. Chen, Varieties of agents in agent-based computational economics: a historical and an interdisciplinary perspective. J. Econ. Dyn. Control **36**(1), 1–25 (2012)

20. S.H. Chen, C.L. Chang, Y.R. Du, Agent-based economic models and econometrics. Knowl. Eng. Rev. **27**(2), 187–219 (2012)

21. S.H. Chen, T.W. Kuo, *Evolutionary Computation in Economics and Finance: A Bibliography* (Physica-Verlag, Heidelberg, 2002), pp. 419–455

22. S.H. Chen, C.H. Yeh, Evolving traders and the business school with genetic programming: a new architecture of the agent-based artificial stock market. J. Econ. Dyn. Control **25**(3), 363–393 (2001). Agent-based Computational Economics (ACE)

23. S.H. Chen, C.H. Yeh, W.C. Lee, Option pricing with genetic programming, in *Genetic Programming 1998: Proceedings of the Third Annual Conference*, ed. by J.R. Koza, W. Banzhaf, K. Chellapilla, K. Deb, M. Dorigo, D.B. Fogel, M.H. Garzon, D.E. Goldberg, H. Iba, R. Riolo (Morgan Kaufmann, Burlington, 1998), pp. 32–37

24. N. Chidambaran, J. Triqueros, C.W.J. Lee, *Option Pricing via Genetic Programming* (Physica-Verlag, Heidelberg, 2002), pp. 383–397

25. I. Contreras, J.I. Hidalgo, L. Nuñez-Letamendía, J.M. Velasco, A meta-grammatical evolutionary process for portfolio selection and trading. Genet. Program. Evol. Mach. **18**(4), 411–431 (2017)

26. J.C. Cox, S.A. Ross, M. Rubinstein, Option pricing: a simplified approach. J. Financ. Econ. **7**(3), 229–263 (1979)

27. S. Cramer, M. Kampouridis, A.A. Freitas, Decomposition genetic programming: an extensive evaluation on rainfall prediction in the context of weather derivatives. Appl. Soft Comput. **70**, 208–224 (2018)

28. S. Cramer, M. Kampouridis, A.A. Freitas, A. Alexandridis, Predicting rainfall in the context of rainfall derivatives using genetic programming, in *2015 IEEE Symposium Series on Computational Intelligence* (2015), pp. 711–718

29. S. Cramer, M. Kampouridis, A.A. Freitas, A. Alexandridis, Pricing rainfall based futures using genetic programming, in *20th European Conference, EvoApplications: European Conference on the Applications of Evolutionary Computation*, vol. 10199 (Springer, Berlin, 2017), pp. 17–33

30. S. Cramer, M. Kampouridis, A.A. Freitas, A. Alexandridis, Stochastic model genetic programming: deriving pricing equations for rainfall weather derivatives. Swarm Evolut. Comput. **46**, 184–200 (2019)

31. S. Cramer, M. Kampouridis, A.A. Freitas, A.K. Alexandridis, An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives. Expert Syst. Appl. **85**, 169–181 (2017)

32. W. Cui, A. Brabazon, M. O'Neill, Evolving dynamic trade execution strategies using grammatical evolution, in *Applications of Evolutionary Computation*, ed. by C. Di Chio (Springer, Berlin, 2010), pp. 192–201

33. W. Cui, A. Brabazon, M. O'Neill, Evolving efficient limit order strategy using grammatical evolution, In *IEEE Congress on Evolutionary Computation* (2010), pp. 1–6

34. W. Cui, A. Brabazon, M. O'Neill, Adaptive trade execution using a grammatical evolution approach. Int. J. Financ. Mark. Deriv. **2**(1/2), 4–31 (2011)

35. G. Deboeck, *Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets* (Wiley, New York, 1994)

36. M.A.H. Dempster, C.M. Jones, A real-time adaptive trading system using genetic programming. Quant. Finance **1**(4), 397–413 (2001)
37. S. Ecca, M. Marchesi, A. Setzu, Modeling and simulation of an artificial stock option market. Comput. Econ. **32**(1), 37–53 (2008)
38. G.P.C. Fung, J.X. Yu, W. Lam, Stock prediction: integrating text mining approach using real-time news. In *2003 IEEE International Conference on Computational Intelligence for Financial Engineering, 2003. Proceedings* (2003), pp. 395–402
39. D.K. Gode, S. Sunder, Allocative efficiency of markets with zero-intelligence traders: market as a partial substitute for individual rationality. J. Polit. Econ. **101**(1), 119–137 (1993)
40. C. Grosan, A. Abraham, Stock market modeling using genetic programming ensembles, in *Genetic Systems Programming: Theory and Experiences*, ed. by N. Nedjah, L.M. Mourelle, A. Abraham (Springer, Berlin, 2006), pp. 131–146. https://doi.org/10.1007/3-540-32498-4_6
41. J. Gypteau, F.E.B. Otero, M. Kampouridis, Generating directional change based trading strategies with genetic programming, in *Applications of Evolutionary Computation*, ed. by A.M. Mora, G. Squillero (Springer, Berlin, 2015), pp. 267–278
42. E. Hemberg, J. Rosen, G. Warner, S. Wijesinghe, U.M. O'Reilly, Tax non-compliance detection using co-evolution of tax evasion risk and audit likelihood, in *Proceedings of the 15th International Conference on Artificial Intelligence and Law, ICAIL '15* (ACM, New York, 2015), pp. 79–88
43. E. Hemberg, J. Rosen, G. Warner, S. Wijesinghe, U.M. O'Reilly, Detecting tax evasion: a co-evolutionary approach. Artif. Intell. Law **24**(2), 149–182 (2016)
44. H. Iba, T. Sasaki, Using genetic programming to predict financial data, in *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406)*, vol. 1 (1999), pp. 244–251
45. K. Izumi, An artificial market model of a foreign exchange market. PhD dissertation, Tokyo University (1999)
46. M.A. Kaboudan, Genetic programming prediction of stock prices. Comput. Econ. **16**(3), 207–236 (2000)
47. M. Kampouridis, A. Alsheddy, E. Tsang, On the investigation of hyper-heuristics on a financial forecasting problem. Ann. Math. Artif. Intell. **68**(4), 225–246 (2013)
48. M. Kampouridis, S.H. Chen, E. Tsang, Market fraction hypothesis: a proposed test. Int. Rev. Financ. Anal. **23**, 41–54 (2012)
49. M. Kampouridis, F.E. Otero, Evolving trading strategies using directional changes. Expert Syst. Appl. **73**, 145–160 (2017)
50. M. Kampouridis, F.E.B. Otero, Heuristic procedures for improving the predictability of a genetic programming financial forecasting algorithm. Soft Comput. **21**(2), 295–310 (2015)
51. M. Kampouridis, E. Tsang, Investment opportunities forecasting: extending the grammar of a GP-based tool. Int. J. Comput. Intell. Syst. **5**(3), 530–541 (2012)
52. M. Kolanovic, R.T. Krishnamachari, Big data and AI strategies: machine learning and alternative data approach to investing. J. P. Morgan Report (2018)
53. J.R. Koza, Genetic programming: a paradigm for genetically breeding populations of computer programs to solve problems. Technical report, Stanford, CA, USA (1990)
54. J.R. Koza, A genetic approach to econometric modeling, in *Economics and Cognitive Science*, ed. by P. Bourgine, B. Walliser (Pergamon Press, Cambridge, 1992), pp. 57–75
55. J.R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection* (MIT Press, Cambridge, 1992)
56. W.B. Langdon, S.M. Gustafson, Genetic programming and evolvable machines: ten years of reviews. Genet. Program. Evol. Mach. **11**(3), 321–338 (2010)
57. F. Larkin, C. Ryan, Good news: using news feeds with genetic programming to predict stock prices, in *Genetic Programming*, ed. by M. O'Neill (Springer, Berlin, 2008), pp. 49–60
58. T. Lensberg, A. Eilifsen, T.E. McKee, Bankruptcy theory development and classification via genetic programming. Eur. J. Oper. Res. **169**(2), 677–697 (2006)
59. S. Martinez-Jaramillo, E.P.K. Tsang, An heterogeneous, endogenous and coevolutionary GP-based financial market. IEEE Trans. Evol. Comput. **13**(1), 33–55 (2009)
60. R.C. Merton, Theory of rational option pricing. Bell J. Econ. Manag. Sci. **4**(1), 141–183 (1973)
61. C. Neely, P. Weller, R. Dittmar, Is technical analysis in the foreign exchange market profitable? A genetic programming approach. J. Financ. Quant. Anal. **32**(4), 405–426 (1997)
62. N.Y. Nikolaev, H. Iba, Genetic programming of polynomial models for financial forecasting, in *Genetic Algorithms and Genetic Programming in Computational Finance, Chap. 5*, ed. by S.H. Chen (Kluwer Academic Press, Dordrecht, 2002), pp. 103–123

63. M. O'Neill, L. Vanneschi, S. Gustafson, W. Banzhaf, Open issues in genetic programming. Genet. Program. Evol. Mach. **11**(3), 339–363 (2010)
64. C.S. Ong, J.J. Huang, G.H. Tzeng, Building credit scoring models using genetic programming. Expert Syst. Appl. **29**(1), 41–47 (2005)
65. S. Salcedo-Sanz, J.L. Fernandez-Villacanas, M.J. Segovia-Varge, C. Bousono-Calzon, Genetic programming for the prediction of insolvency in non-life insurance companies. Comput. Oper. Res. **32**(4), 749–765 (2005)
66. A. Samitas, S. Polyzos, C. Siriopoulos, Brexit and financial stability: an agent-based simulation. Econ. Model. **69**, 181–192 (2018)
67. H. Schmidbauer, A. Rösch, T. Sezer, V.S. Tunalioğlu, Robust trading rule selection and forecasting accuracy. J. Syst. Sci. Complex. **27**(1), 169–180 (2014)
68. K. Sörensen, Metaheuristics—the metaphor exposed. Int. Trans. Oper. Res. **22**(1), 3–18 (2015)
69. E. Tsang, J. Chen, Regime change detection using directional change indicators in the foreign exchange market to chart brexit. IEEE Trans. Emerg. Topics Comput. Intell. **2**(3), 185–193 (2018)
70. E.P.K. Tsang, R. Tao, A. Serguieva, S. Ma, Profiling high-frequency equity price movements in directional changes. Quant. Finance **17**(2), 217–225 (2017). https://doi.org/10.1080/14697688.2016.1164887
71. C. Tuite, M. O'Neill, A. Brabazon, Economic and financial modeling with genetic programming, in *The Oxford Handbook of Computational Economics and Finance, Chapter 8*, ed. by S.H. Chen, M. Kaboudan, Y.R. Du (Oxford Handbooks Online, Oxford, 2018), pp. 267–289
72. N. Wagner, Z. Michalewicz, M. Khouja, R.R. McGregor, Time series forecasting for dynamic environments: the dyfor genetic program model. IEEE Trans. Evol. Comput. **11**(4), 433–452 (2007)
73. P.A. Whigham, R. Withanawasam, Evolving a robust trader in a cyclic double auction market, in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation, GECCO '11* (ACM, New York, 2011), pp. 1451–1458
74. H. White, A reality check for data snooping. Econometrica **68**(5), 1097–1126 (2000)
75. W. Yan, C.D. Clack, Evolving robust GP solutions for hedge fund stock selection in emerging markets. Soft. Comput. **15**(1), 37–50 (2011)
76. Z. Yin, A. Brabazon, C. O'Sullivan, Adaptive genetic programming for option pricing, in *Proceedings of the 9th Annual Conference Companion on Genetic and Evolutionary Computation, GECCO '07* (ACM, New York, 2007), pp. 2588–2594
77. Z. Yin, A. Brabazon, C. O'Sullivan, P.A. Hamill, A genetic programming approach for delta hedging. Genet. Program. Evol. Mach. **20**(1), 67–92 (2019)
78. Z. Yin, A. Brabazon, C. O'Sullivan, M. O'Neill, A genetic programming approach for delta hedging, in *2015 IEEE Congress on Evolutionary Computation (CEC)* (2015), pp. 3312–3318