International Journal of Forecasting 33 (2017) 21–47

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

A comparison of wavelet networks and genetic programming in the context of temperature derivatives

Antonis K. Alexandridis[a,*], Michael Kampouridis[b], Sam Cramer[b]

[a] School of Mathematics, Statistics and Actuarial Science, University of Kent, United Kingdom
[b] School of Computing, University of Kent, United Kingdom

ARTICLE INFO

Keywords:
Weather derivatives
Wavelet networks
Temperature derivatives
Genetic programming
Modelling
Forecasting

ABSTRACT

The purpose of this study is to develop a model that describes the dynamics of the daily average temperature accurately in the context of weather derivatives pricing. More precisely, we compare two state-of-the-art machine learning algorithms, namely wavelet networks and genetic programming, with the classic linear approaches that are used widely in the pricing of temperature derivatives in the financial weather market, as well as with various machine learning benchmark models such as neural networks, radial basis functions and support vector regression. The accuracy of the valuation process depends on the accuracy of the temperature forecasts. Our proposed models are evaluated and compared, both in-sample and out-of-sample, in various locations where weather derivatives are traded. Furthermore, we expand our analysis by examining the stability of the forecasting models relative to the forecasting horizon. Our findings suggest that the proposed nonlinear methods outperform the alternative linear models significantly, with wavelet networks ranking first, and that they can be used for accurate weather derivative pricing in the weather market.

1. Introduction

This paper uses wavelet networks (WNs) and genetic programming (GP) to describe the dynamics of the daily average temperature (DAT), in the context of weather derivatives pricing. The proposed methods are evaluated both in-sample and out-of-sample against various linear and non-linear models that have been proposed in the literature.

Recently, a new class of financial instruments, known as "weather derivatives" has been introduced. Weather derivatives are financial instruments that can be used by organizations or individuals to reduce the risk associated with adverse or unexpected weather conditions, as part of a risk management strategy (Alexandridis & Zapranis, 2013a). Just like traditional contingent claims, the payoffs of which depend upon the price of some fundamental, a weather derivative has an underlying measure such as rainfall, temperature, humidity, or snowfall. However, they differ from other derivatives in that the underlying asset has no value and cannot be stored or traded, but at the same time must be quantified in order to be introduced in the weather derivative. To do this, temperature, rainfall, precipitation, or snowfall indices are introduced as underlying assets. However, the majority of the weather derivatives have a temperature index as the underlying asset. Hence, this study focuses only on temperature derivatives.

Studies have shown that about $1 trillion of the US economy is exposed directly to weather risk (Challis, 1999;

* Corresponding author.
E-mail address: A.Alexandridis@kent.ac.uk (A.K. Alexandridis).

http://dx.doi.org/10.1016/j.ijforecast.2016.07.002

Hanley, 1999). Today, weather derivatives are used for hedging purposes by companies and industries whose profits can be affected adversely by unseasonal weather, and for speculative purposes by hedge funds and others who are interested in capitalising on these volatile markets. Weather derivatives are used to hedge volume risk, rather than price risk.

It is essential to have a model that (i) describes the temperature dynamics accurately, (ii) describes the evolution of the temperature accurately, and (iii) can be used to derive closed form solutions for the pricing of temperature derivatives. In complete markets, the cash flows of any strategy can be replicated by a synthetic one. In contrast, the weather market is an incomplete market, in the sense that the underlying asset has no value and cannot be stored, and hence, no replicating portfolio can be constructed. Thus, modelling and pricing the weather market are challenging issues. In this paper, we focus on the problem of temperature modelling. It is of paramount importance to address this problem before doing any investigation into the actual pricing of the derivatives.

There has been quite a significant amount of work done to date in the area of modelling the temperature over a certain time period. Early studies tried to model different temperature indices directly, such as heating degree days (HDD) or the cumulative average temperature (CAT).[1] Following this path, a model is formulated so as to describe the statistical properties of the corresponding index (Davis, 2001; Dorfleitner & Wimmer, 2010; Geman & Leonardi, 2005; Jewson, Brix, & Ziehmann, 2005). One obvious drawback of this approach is that a different model must be used for each index when formulating the temperature index, such as HDD, as a normal or lognormal process, meaning that a lot of information both in common and extreme events is lost; e.g., HDD is bounded by zero (Alexandridis & Zapranis, 2013a).

More recent studies have utilized dynamic models, which simulate the future behavior of DAT directly. The estimated dynamic models can be used to derive the corresponding indices and price various temperature derivatives (Alexandridis & Zapranis, 2013a). In principle, using models for daily temperatures can lead to more accurate pricing than modelling temperature indices. The continuous processes used for modeling DAT usually take a mean-reverting form, which has to be discretized in order to estimate its various parameters.

Most models can be written as nested forms of a mean-reverting Ornstein–Uhlenbeck (O–U) process. Alaton, Djehince, and Stillberg (2002) propose the use of an O–U model with seasonalities in the mean, using a sinusoidal function and a linear trend in order to capture urbanization and climate changes. Similarly, Benth and Saltyte-Benth (2007) use truncated Fourier series in order to capture the seasonality in the mean and volatility. In a more recent paper, Benth, Saltyte-Benth, and Koekebakker (2007) propose the use of a continuous autoregressive model. Using 40 years of data in Stockholm, their results indicate that their proposed framework is sufficient to explain the autoregressive temperature dynamics. Overall, the fit is very good; however, the normality hypothesis is rejected even though the distribution of the residuals is close to normal.

A common denominator in all of the works mentioned above is that they use linear models, such as autoregressive moving average models (ARMA) or their continuous equivalents (Benth & Saltyte-Benth, 2007). However, a fundamental problem of such models is the assumption of linearity, which cannot capture some features that occur commonly in real-world data, such as asymmetric cycles and outliers (Agapitos, OŃeill, & Brabazon, 2012b). On the other hand, nonlinear models can encapsulate the time dependency of the dynamics of the temperature evolution, and can provide a much better fit to the temperature data than the classic linear alternatives.

One example of a nonlinear work is that by Zapranis and Alexandridis (2008), who used nonlinear non-parametric neural networks (NNs) to capture the daily variations of the speed at which the temperature reverts to its seasonal mean. Their results indicated that they had managed to isolate the Gaussian factor in the residuals, which is crucial for accurate pricing. Zapranis and Alexandridis (2009) used NNs to model the seasonal component of the residual variance of a mean-reverting O–U temperature process, with seasonality in the level and volatility. They validated their proposed method on more than 100 years of data collected from Paris, and their results showed a significant improvement over more traditional alternatives, regarding the statistical properties of the temperature process. This is important, since small misspecifications in the temperature process can lead to large pricing errors. However, although the distributional statistics were improved significantly, the normality assumption of the residuals was rejected.

NNs have the ability to approximate any deterministic nonlinear process, with little knowledge and no assumptions regarding the nature of the process. However, the classical sigmoid NNs have a series of drawbacks. Typically, the initial values of the NN's weights are chosen randomly, which is generally accompanied by extended training times. In addition, when the transfer function is of sigmoidal type, there is always a significant chance that the training algorithm will converge to a local minimum. Finally, there is no theoretical link between the specific parametrization of a sigmoidal activation function and the optimal network architecture, i.e., model complexity.

In this paper, we continue to look into nonlinear models, but we move away from neural networks. Instead, we look into two other algorithms from the field of machine learning (Mitchell, 1997): wavelet networks (WNs) and genetic programming (GP). The two proposed nonlinear methods will then be used to model the DAT. There are various reasons why we focus on these two nonlinear models. First, we want to avoid the black-boxes produced by alternative nonlinear models, such as NNs and support vector machines (SVM). Second, both models have many desirable properties, as it is explained below.

One of the main advantages of GP is its ability to produce white-box (interpretable) models, which allows traders to visualise the candidate solutions, and thus the

[1] The CAT and HDD indices are explained in Section 2.