Annals of Operations Research
https://doi.org/10.1007/s10479-023-05483-3

ORIGINAL RESEARCH

Minimizing the impact of geographical basis risk on weather derivatives

Mina D'Aversa[1] · Alessandra Mainini[2] · Enrico Moretto[3] · Silvana Stefani[2] · Pierpaolo Uberti[4]

Accepted: 20 June 2023

Abstract

In the last decade, the index-based weather products (also called weather derivatives) have been gaining attention in the climate resilience discussion. Weather derivatives are designed to help companies hedging against climate variability. These products, that can be market-traded or over-the-counter, compensate individuals based on a pre-defined weather index. Thus, pay-offs of a weather derivative depend on a weather index and not, as with traditional types of insurance, on the actual amount of money lost due to adverse weather. One of the major drawbacks that may prevent weather derivatives to catch on is the impact of the Geographical Basis Risk (GBR), that is the deviation of weather conditions at different locations. In fact, when the reference weather station is not located in the immediate vicinity of the site of interest the hedging effectiveness may be reduced. In this paper, we contribute to the existing literature on GBR by proposing an optimization method that may help in offering a tailored solution, while at the same time keeping a standardized instrument as a reference. Using a historical record of Italian temperatures, strikes for temperatures are the choice variables of a penalty function containing pay-offs of a reference station and all other

✉ Enrico Moretto
enrico.moretto@unimib.it

Mina D'Aversa
m.daversa@campus.unimib.it

Alessandra Mainini
alessandra.mainini@unicatt.it

Silvana Stefani
silvana.stefani@unicatt.it

Pierpaolo Uberti
pierpaolo.uberti@unimib.it

1 Università di Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italia

2 Dipartimento di Scienze Economiche e Sociali, Università Cattolica del Sacro Cuore, Via Emilia Parmense 84, 29122 Piacenza, Italia

3 Dipartimento di Economia, Metodi Quantitativi e Strategie d'Impresa, Università di Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italia

4 Dipartimento di Statistica e Metodi Quantitativi, Università di Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italia

Published online: 14 July 2023

stations. Further, altitude and latitude of meteorological stations are shown to be relevant predictors to explain GBR. This can be an interesting starting point for the design of weather derivatives, since, from a unique station where the "reference" derivative is priced, all the other stations may be easily settled.

Keywords Geographical basis risk · Weather derivatives · Hedging effectiveness · Climate risk

1 Introduction

Since the beginning of the 21st century, climate topics are gaining interest at a very accelerated pace. One major consequence is the accrued usage of an index-based weather insurance which compensates agents who buy protection against meteorological adverse events. Unlike traditional insurance contracts, these financial instruments can be seen as derivative contracts whose advantages range from greater time and cost effectiveness to moral hazard mitigation. Such instruments are commonly known as index-based weather products or Weather Derivatives (WDs). They are either traded on a regulated market or settled over the counter (OTC).

Consider, for instance, a farmer who deals with droughts, floods, and extreme temperatures (see Bucheli et al. 2021 for an analysis of tailored crop and climate change insurances) or a touristic activity whose profits rely on abundant snowfall in the winter season or on a rainy summertime (see Franzoni and Pelizzari 2021). Similarly to standard option and future contracts, WDs can help managing weather related risks.

A large portion of financial literature covering this topic deals with WDs' quantitative valuation. For a comprehensive guide about this issue, see Jewson and Brix (2005) and, more recently, Stefani et al. (2018; 2020).

Setting aside the valuation problem, several challenges are to be considered in a WDs context. On one side, these financial instruments reduce the need for insurance contracts, as meteorological risks can be differently hedged. This is important as climate changes are implying a greater risk of larger claims both in terms of frequency and severity of weather events (see Hellmuth et al. 2009 and, more recently, Abdi et al. 2022). However, WDs are highly standardized contracts: if this makes their market liquid, these contracts cannot be easily adapted to specific needs of local areas. This explains WDs' limited application.

Such instruments are actively traded on the Chicago Mercantile Exchange[1] (CME). With regards to temperature-based instruments,[2] as of 2022 these contracts are HDD (Heating Degree Days), CDD (Cooling Degree Days), and CAT (Cumulative Average Temperature). Pay-offs for HDDs and CDDs resemble European options' ones, with $k = 18°\mathrm{C}$ as strike price. Unlike derivatives written on stocks or commodities, here the underlying is not a negotiable asset.

This contingent claim approach is exploited not only for strictly financial reasons. A page available on the Eurostat website[3] exploits CDDs and HDDs to determine the energy requirement for cooling and heating buildings across Europe.

[1] https://www.cmegroup.com/education/articles-and-reports/weather-options-overview.html - website accessed on December 17th, 2022.

[2] The underlying temperatures are measured with daily frequency in a number of reference cities (nine in the US as well as London, Amsterdam, and Tokyo).

[3] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Heating_and_cooling_degree_days_-_statistics#Heating_and_cooling_degree_days_at_EU_level - website accessed on December 17th, 2022.

In Musshoff et al. (2011), the authors analyze WDs as an effective tool for the protection of farmers against volumetric risk, that is the fact that crops production might be less than expected. Using real yield and weather data from Northeastern Germany, the authors conclude that the use of rainfall options has a considerable hedging effect when the reference weather station is located in the immediate vicinity of the site of production and when there is an evident relationship between yield and rainfall. WDs could therefore generate support for farming activities that operate under risky conditions.

A new temperature index that encompasses a long-term memory effect and that can represent the underlying of a WD is proposed in Castellano et al. (2020). This approach might be of use to hedge volumetric risk.

The valuation of a CDD contract based on Malaysian temperatures is analyzed in Taib and Benth (2012). Being a developing country, Malaysia relies heavily on its agricultural production. Any situation that may harm its crops should then be effectively hedged. The profit/loss distribution generated by the contract is investigated in the perspective of both counterparts.

In Franzoni and Pelizzari (2021), WDs with precipitation as underlying variable are seen with a touristic business perspective. Rain has an obvious negative impact on the performance of hospitality industry both in the short and medium runs. In this contribution, the authors design a rainfall option capable of compensating for the lack of revenues caused by an excess of precipitation, and price it in a risk-neutral and multidimensional framework using the Monte Carlo method. Their main contribution is, along with the valuation of such WDs, the determination of the optimal number of contracts capable of minimizing risk due to rainfall.

Going back to the main goal of this contribution, from both a theoretical and practical perspectives, the relevant issue is two-fold: on one side, standardization is needed for the existence of a sufficiently liquid market, as in Musshoff et al. (2011); this leads to the determination of fair prices. On the other one, though, it is evident that WDs are poorly effective in dealing with meteorological risk in places that are far from the reference location where meteorological data is collected. Consider, for instance, an agricultural activity in southern Europe with respect to a WD written on Amsterdam's temperatures. As a consequence, this risk remains with the protection-seeking side of the derivative, carrying to an incomplete and therefore rather ineffective hedge. This is what goes under the name of Geographical Basis Risk (GBR): a full description of this notion can be found in Alexandridis and Zapranis (2012). Formally stated, GBR is the deviation of weather conditions at different locations, and the consequent interference with WDs' hedging effectiveness.

In Gyamerah et al. (2019), GBR is tackled using a spatial-temporal pricing model and exploiting the risk-neutral approach to derivative pricing. The standard model in this instance is the Ornstein-Uhlenbeck process (see Benth and Benth 2007) with a regime-switching addendum that encompasses periods in which extreme temperatures replace the historical ones. The spatial component of this approach is modelled using a weighted basket of standard WDs.

An analysis of the impact of rainfall WDs in the agricultural sector is performed in East (2005). Meteorological data are amalgamated so that WDs can be tailored with a sufficiently accurate degree of freedom for specific locations, offering another way of dealing with GBR. The main result of this contribution is that these contracts not only are effective under a theoretical point of view to mitigate meteorological risk but, from a practical perspective, are perceived by farmers as an hedging instrument worth acquiring.

In Ritter et al. (2014), it is shown that GBR can be reduced by regional diversification, and the pay-off of a hypothetical derivative can be approximated by the weighted combination of pay-offs measured in this location's neighbourhood. The authors conclude that the error decreases about 20% when using more than one neighbouring locations.

This being said, the aim of this article is to provide a different methodology capable of dealing with GBR. The approach is model-free: pay-offs of CDDs and HDD for a reference measurement station are computed at first using observed data. Strikes for all other stations are the choice variable of an optimization problem in which we minimize the square difference in pay-offs with respect to those of the reference one. Then, in-sample and out-of-sample analysis are performed, showing good results validated by regressions output. The main result is that a tailored version of WDs for each measurement station can be easily obtained by fixing an appropriate strike price. This allows to price fairly these derivatives, letting that, even in a wide area, a single WD can provide a reliable hedging that is valid for all other locations.

There is no need to say that goodness of meteorological data is of utter importance. There is a number of datasets that can be exploited. Still, in many cases historical observations can contain errors or being missed. A meteorological station could also be disrupted and newer ones added. Frequently, then, data ends up being not updated and time series on multiple locations have different lengths. In the database [4] exploited here, Italian temperatures are measured and collected from 1973 to 2022 for 197 weather stations located all over the country. Further, for each meteorological station (identified by an ID Code), latitude lat_i, longitude long_i, and altitude above sea level alt_i, $i = 1, \ldots, 197$ are available.

Rome is taken as the reference station. It results that, once the optimal strike prices are computed, the height and latitude of measurement stations turn out being an effective predictor of these strike prices. On the other hand, distance from the reference station to the other locations is not relevant.

Unlike other contributions, this approach is easy to implement and to be explained to potential buyers of risk mitigating contracts. Here the innovation is that strike prices, considered so far fixed, become now the key variable for adapting pay-offs to specific locations.

This paper is structured as follows: Sect. 2 provides the definition of Weather Derivatives. Section 3 explains how GBR can be hedged. Section 4 provides the in- and out-of-sample analysis. Section 5 concludes.

2 Definition of WDs

A Weather Derivative (WD) is a financial contract whose pay-off depends on the behaviour of some meteorological underlying variable such as temperature, wind, and rainfall.

This paper focuses on a temperature-based WD: the underlying variable is the daily arithmetic mean between the maximum $T_{d,i}^{\max}$ and minimum $T_{d,i}^{\min}$ observed temperatures for the $i-$th meteorological station ($i = 1, ..., n$) at day d ($d = 1, \ldots, D$), which is denoted with

$$\bar{T}_{d,i} = \frac{T_{d,i}^{\max} + T_{d,i}^{\min}}{2}.$$

The two most common contingent claims dealing with temperature are the Cooling Degree Days (CDD) and Heating Degree Days (HDD) contracts. Their pay-offs are the sum, over a period of D days, of cash flows determined on a daily basis. In particular, the pay-off of a CDD contract is

[4] This database is not of public domain and is maintained by Luigi Mariani and Franco Zavatti.

$$\mathrm{CDD}_i = \lambda \sum_{d=1}^{D} \max(\bar{T}_{d,i} - k_i; 0) \tag{1}$$

where λ is the 'tick size', i.e. the amount of money that translates temperatures into cash flows. Here, for sake of simplicity, $\lambda = 1$. This derivative rewards its subscriber whenever temperatures go beyond the reference temperature k_i and, in most cases, its time interval encompasses summer months.

On the contrary, a HDD contract rewards its subscriber against temperatures going below the reference temperature k_i; its cumulative pay-off is defined as

$$\mathrm{HDD}_i = \lambda \sum_{d=1}^{D} \max(k_i - \bar{T}_{d,i}; 0). \tag{2}$$

Here, the time range covers winter months as, in the northern hemisphere, December, January, and February.

CDDs and HDDs are traded on the Chicago Mercantile Exchange (CME) (cf. Alexandridis and Zapranis 2012) where strikes temperatures[5] are equal to $k = 18°$. According to the European Environment Agency,[6] different thresholds might be considered: $k = 15.5°$ for HDDs and $k = 22°$ for CDDs. In what follows, for the reference meteorological station (REF) we will choose $k_{REF} = 18°$ as strike.

3 A model-free approach to GBR

In this Section we present the theoretical contribution of this article: we deal with GBR, which is, as said in the Introduction, an important drawback that limits the diffusion of WDs. Unlike standard approaches, where WDs are priced using some model that determines the underlying behaviour, here no assumption is made on such dynamics. Once a reference station and a given strike price are chosen, the relevant variable becomes the strike price of all other stations. These values are numerically computed in order to obtain pay-offs that are as similar as possible to the one of the reference station. Large geographical areas (in our case the whole territory of Italy) might then be covered.

3.1 The optimization problem

Suppose that a standard HDD/CDD derivative, having as underlying the temperature recorded in a REF station and $k_{REF} = 18°$ as strike price, is traded on a market.

The rationale behind is: if an hypothetical CDD/HDD contract is quoted, having REF as reference station, and if a contract for another weather station can be replicated with good approximation in terms of REF's pay-off, an economic agent may be willing to buy, for hedging purposes, an equivalent product quoted on the basis of the local data recorded.

For each meteorological station, the optimal strike price k_i^* is calculated so that the daily pay-offs generated in the $i-$th station become as close as possible, in a square error minimization sense, to those of the REF station.

[5] All temperatures are in Celsius scale.

[6] https://www.eea.europa.eu/data-and-maps/indicators/heating-degree-days-2 - website accessed on December 17th, 2022.

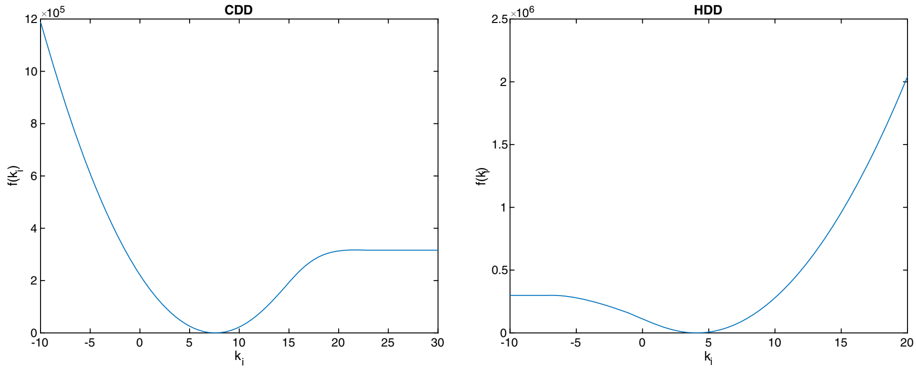

Fig. 1 Shape of the objective function of problem (3) (left plot) and problem (4) (right plot); Year 2021. See second row (Brennero Pass station - $i = 2$) in Table 1

Namely, for each $i = 1, ..., n, i \neq REF$,

$$k_i^* = \arg\min_{k_i} \sum_{d=1}^{D} [\max(\bar{T}_{d,REF} - 18; 0) - \max(\bar{T}_{d,i} - k_i; 0)]^2 \tag{3}$$

in the case of a CDD contract and

$$k_i^* = \arg\min_{k_i} \sum_{d=1}^{D} [\max(18 - \bar{T}_{d,REF}; 0) - \max(k_i - \bar{T}_{d,i}; 0)]^2 \tag{4}$$

for the HDD one.

The analysis concentrates at first on a CDD contract with the months of July and August ($D = 62$) as its time span. See Table 1, first row, for the cumulated pay-off of Rome.

The second contract is an HDD, in the time range of December, January, February ($D = 90$ or 91 according to the years), with, again, Rome as the reference station. The cumulated pay-off of a standard HDD contract is generated.

The analysis then proceeds with the solution of optimization problems (3) and (4). Figure 1 depicts the objective functions for temperatures measured in 2021 at the Brennero Pass station.

The lack of convexity in the domain of these function and the presence of a flat region make the problem non-trivial. Its solution deserves an *ad-hoc* algorithm. The one exploited here was presented in Figini and Uberti (2010).

It is now interesting to ascertain the impact of GBR in Italy. Figure 2 plots pay-offs computed according to (1), with strike $k = 18°$ in year 2021, for all weather stations. The Figure confirms the effect of the distance on contracts where the strike price is kept fixed for all stations. Pay-offs range between 0 and approximately 700 euros, while the REF pay-off is 559.25 (see Table 1, first row).

4 Results

This Section presents the main numerical results of this article and shows the in-sample and out-of-sample analysis.

Table 1 Altitude, longitude, latitude, optimal strike, and optimal pay-off for a sample of 32 weather stations in year 2021

(ID-Location)	Alt	Long	Lat	k_i^*	Pay-off
(i = 1) 162400-Rome	95	12480	41900	18.00	559.25
(i = 2) 160150-Brennero Pass	1362	11500	46983	7.59	524.92
(i = 3) 160140-Vipiteno	921	11430	46883	9.91	524.53
(i = 4) 160083-Resia Pass	1800	10500	46833	6.00	530.56
(i = 5) 160080-S.Valentino alla Muta	1461	10530	46750	6.06	529.59
(i = 6) 160330-Dobbiaco	1226	12210	46733	7.92	523.54
(i = 7) 160400-Tarvisio	778	13580	46500	9.38	532.98
(i = 8) 160410-Tarvisio	778	13580	46500	11.53	555.78
(i = 9) 160200-Bolzano	240	11320	46460	14.29	523.69
(i = 10) 160210-Rolle Pass	2006	11780	46300	3.91	520.83
(i = 11) 160220-Paganella Mountain	2129	11030	46150	2.20	526.25
(i = 12) 160370-Aviano (USAF)	125	12610	46033	14.99	541.17
(i = 13) 160440-Udine/Campoformido	94	13180	46033	14.57	550.47
(i = 14) 160360-Aviano AB	117	12590	46032	15.02	540.73
(i = 15) 692894-Aviano AFB/TEST	127	12360	46030	15.22	537.20
(i = 16) 160365-Aviano (USAF)	126	12610	46017	14.97	541.60
(i = 17) 160450-Rivolto	54	13040	45979	14.98	554.34
(i = 18) 160700-Grigna Settentrion	2403	9380	45950	3.94	529.52
(i = 19) 160455-Gorizia	63	13630	45950	14.90	551.35
(i = 20) 160520-Pian Rosa (MTN TOP)	3488	7700	45933	−6.41	523.18
(i = 21) 160720-Bisbino Mountain	1322	9060	45867	9.25	526.32
(i = 22) 160920-Grappa Mountain	1775	11800	45867	7.15	526.58
(i = 23) 161080-Ronchi dei Legionari	11	13470	45828	15.03	544.38
(i = 24) 161070-Concordia Sagittaria	2	12850	45750	15.13	547.93
(i = 25) 160540-Aosta Pollein	551	7350	45733	13.01	567.66
(i = 26) 160980-Istrana	41	12080	45685	16.74	523.70
(i = 27) 160760-Bergamo Orio al Serio	238	9700	45674	14.45	528.67
(i = 28) 161100-Trieste	20	13750	45650	16.14	574.83
(i = 29) 160990-Treviso	18	12190	45648	16.11	528.17
(i = 30) 160660-Malpensa	233	8720	45631	14.72	520.03
(i = 31) 160940-Vicenza	39	11530	45573	16.07	532.90
(i = 32) 160640-Cameri	178	8660	45530	14.90	524.75

4.1 In-sample analysis

The solution of optimization problems (3) and (4), (i.e. the optimal strike price) allows to compute, for each station and year by year, a pay-off replicating that of the reference station.

Table 1 displays a sample of optimal strikes and pay-offs for 32 weather stations in year 2021. Rome, the reference station located at 95 m above sea level whose pay-off is 559.29 euro computed with strike $k = 18°$, is the first entry.

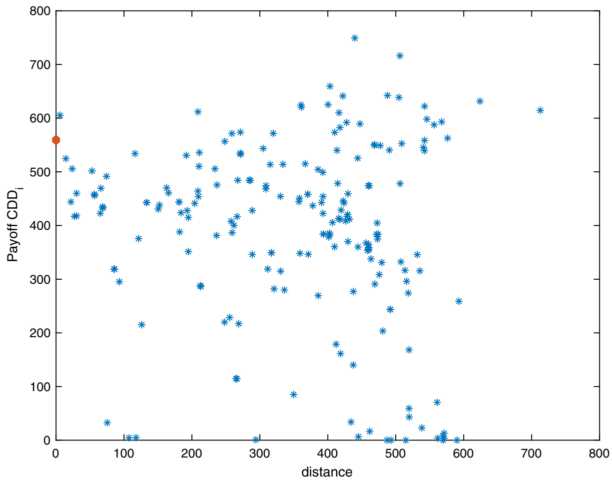

Fig. 2 Pay-offs of CDD contracts vs distance (measured in kilometers) from the REF station (red dot) for 197 Italian meteorological stations with $k = 18°$-Year 2021

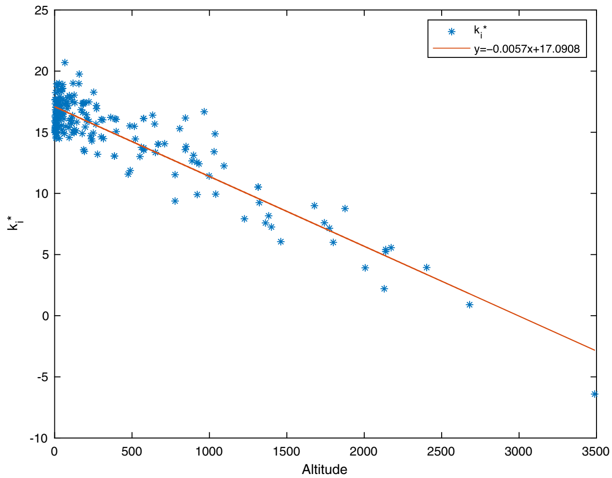

Fig. 3 Scatter plot and linear regression between altitude (x axis) and optimal strike price (y axis) for the 197 weather stations-Year 2021

The next step is to investigate the determinants of k_i^*. At first, the relationship existing between alt_i and k_i^* is studied. Figure 3 represents, for a CDD contract, the scatter plot for all the meteorological stations in 2021: intercept (17.0908) and slope (−0.0057) are both significant (confidence level 99%). Further, $R^2 = 0.83$. For weather stations located at an altitude close to the sea level, strikes are in a range close to k_{REF}. As the altitude increases, optimal strikes decrease, justifying the negative slope. Results for the other years are very similar and are resumed in Table 2.

In Table 2, intercepts, slopes and R^2 for a linear regression with height as independent variable and optimal strike as the dependent one are reported for all years in the dataset. It is worth remaining that all R^2 coefficients are close to 1, varying in the range $[0.78 - 0.92]$.

Table 2 Regression analysis-altitude vs optimal strike (CDD contracts) for all meteorological stations; years 1973-2022

Year	Slope	Intercept	R^2	Year	Slope	Intercept	R^2
1973	−0.0056	17.8136	0.8914	1998	−0.0055	17.5895	0.8964
1974	−0.0054	17.6381	0.9166	1999	−0.0056	17.9345	0.9132
1975	−0.0059	17.7588	0.9217	2000	−0.0058	18.3914	0.8025
1976	−0.0059	17.9281	0.9013	2001	−0.0054	18.9497	0.8369
1977	−0.0060	18.7725	0.8533	2002	−0.0054	18.6391	0.8686
1978	−0.0058	18.3574	0.8929	2003	−0.0055	18.8011	0.8587
1979	−0.0059	17.8124	0.8870	2004	−0.0054	18.8497	0.8738
1980	−0.0056	18.0332	0.8985	2005	−0.0053	18.9013	0.8500
1981	−0.0056	18.0615	0.9005	2006	−0.0057	18.9049	0.8556
1982	−0.0057	17.8232	0.8884	2007	−0.0059	19.0706	0.7841
1983	−0.0055	18.1493	0.9114	2008	−0.0058	19.4129	0.7899
1984	−0.0056	18.2992	0.9035	2009	−0.0055	19.7853	0.8237
1985	−0.0056	17.7762	0.9079	2010	−0.0051	18.2973	0.8558
1986	−0.0057	17.6812	0.9136	2011	−0.0054	18.2365	0.7834
1987	−0.0058	17.3962	0.8514	2012	−0.0053	18.5646	0.8164
1988	−0.0056	17.7792	0.8642	2013	−0.0053	18.3480	0.8767
1989	−0.0055	18.0785	0.8383	2014	−0.0056	18.2828	0.8717
1990	−0.0051	17.9621	0.8170	2015	−0.0056	18.2570	0.8531
1991	−0.0051	18.0684	0.8553	2016	−0.0056	17.8659	0.8606
1992	−0.0053	17.4434	0.9006	2017	−0.0055	17.2242	0.8612
1993	−0.0055	17.2130	0.8162	2018	−0.0056	17.4090	0.8409
1994	−0.0054	17.3202	0.8462	2019	−0.0056	17.5887	0.9176
1995	−0.0056	17.4441	0.8938	2020	−0.0057	17.1247	0.8819
1996	−0.0057	17.5010	0.8725	2021	−0.0057	17.0908	0.8363
1997	−0.0057	17.4904	0.8497	2022	−0.0057	17.2539	0.8102

All parameters significant at a 99% level

These outcomes provide a strong evidence supporting the stability through time of the relation between height and optimal strikes. Such a stability is a necessary condition to use historical data in predicting the future k_i^* in the out-of-sample framework (see the almost constant behaviour, with respect to time, of both slope and intercept in Table 2).

On the other hand, the geographical distance between the REF station and all the others does not appear to be relevant. Indeed, Table 3 reports values for R^2 in the range [0.0030 – 0.0437], showing a very poor explanatory performance of the geographical distance of each station to the REF one with respect to the optimal strikes. This is confirmed by the regression analysis displayed in Fig. 4, where the linear relationship between distances from the reference station to all other stations and k_i^* is reported. The plot shows a scattered behaviour, possibly driven by the coexistence of two regimes: a stationary one (see stations with a distance less than 400) and another with a steep descending behaviour (see stations with distances larger than 400).

Table 3 Regression analysis-distance from the REF station versus optimal strike k_i^* (1973-2022)

Year	Slope	Intercept	R^2	Year	Slope	Intercept	R^2
1973	−0.0031	16.6757*	0.0191	1998	−0.0039	16.7615*	0.0312
1974	−0.0025	16.3614*	0.0139	1999	−0.0032	16.8374*	0.0210
1975	−0.0032	16.5257*	0.0186	2000	−0.0044	17.6223*	0.0323
1976	−0.0013	16.0136*	0.0030	2001	−0.0033	17.9539*	0.0216
1977	−0.0030	17.4218*	0.0146	2002	−0.0027	17.4193*	0.0149
1978	−0.0030	17.0984*	0.0168	2003	−0.0026	17.5000*	0.0129
1979	−0.0037	16.7458*	0.0240	2004	−0.0039	18.0740*	0.0322
1980	−0.0027	16.7336*	0.0144	2005	−0.0037	18.0854*	0.0291
1981	−0.0030	16.8957*	0.0185	2006	−0.0031	17.7449*	0.0183
1982	−0.0032	16.6741*	0.0193	2007	−0.0039	18.0990*	0.0242
1983	−0.0030	16.9966*	0.0185	2008	−0.0043	18.5759*	0.0293
1984	−0.0018	16.6871*	0.0064	2009	−0.0040*	18.9829*	0.0301
1985	−0.0029	16.5456*	0.0169	2010	−0.0040	17.6356*	0.0354
1986	−0.0030	16.4562*	0.0176	2011	−0.0040*	17.4783*	0.0300
1987	−0.0041	16.5065*	0.0291	2012	−0.0043	17.9623*	0.0377
1988	−0.0035	16.7667*	0.0234	2013	−0.0036	17.4868*	0.0277
1989	−0.0033	17.0439*	0.0211	2014	−0.0035	17.2974*	0.0243
1990	−0.0034	17.1240*	0.0253	2015	−0.0038	17.3382*	0.0268
1991	−0.0029	17.0373*	0.0191	2016	−0.0046*	17.2210*	0.0395
1992	−0.0031	16.4080*	0.0215	2017	−0.0047*	16.6759*	0.0437
1993	−0.0040	16.4471*	0.0310	2018	−0.0047*	16.8274*	0.0415
1994	−0.0041	16.5730*	0.0329	2019	−0.0036	16.6131*	0.0265
1995	−0.0032	16.3154*	0.0197	2020	−0.0046*	16.4793 *	0.0406
1996	−0.0025	16.0691*	0.0111	2021	−0.0046*	16.4167*	0.0375
1997	−0.0027	16.1622*	0.0131	2022	−0.0040	16.3755*	0.0278

Parameters significant at 99% are denoted with *

A more accurate fitting for k_i^* is displayed in Table 4. Here, the independent variables are height and latitude. The regression shows that all estimates are significant at 99% level. Further, all R^2 vary in the $[0.87 - 0.96]$ range.

To resume for a CDD contract optimal strike for each station in the Italian territory is well explained by altitude and latitude.

The same analysis, carried on for a HDD contract, gives a reduced goodness of fit using altitude alone as independent variable (R^2 ranges in interval $[0.45 - 0.67]$). On the contrary, introducing also latitude yields large values for R^2, in the $[0.83 - 0.94]$ range: see Tables 5 and 6.

Before concluding this sub-section, it is worth pointing out that the choice of the REF station is totally arbitrary and does not affect the final results. In fact, this is one of the relevant features of our findings: the REF station could be chosen independently of the geographical position of the hedging-buying agent. The fact that WDs written on REF's temperatures are traded on a market will also guarantee adequate pricing.

Table 4 Yearly multiple linear regressions between altitude, latitude, and optimal strike (1973-2022)-CDD contracts

Year	alt	lat	Intercept	R^2	Year	alt	lat	Intercept	R^2
1973	−0.3221	−0.0052	31.4282	0.9491	1998	−0.2701	−0.0052	29.0044	0.9385
1974	−0.1861	−0.0052	25.5049	0.9374	1999	−0.2171	−0.0053	27.1114	0.9448
1975	−0.2634	−0.0056	28.8914	0.9576	2000	−0.5559	−0.0052	41.8863	0.9448
1976	−0.3154	−0.0056	31.2587	0.9503	2001	−0.3682	−0.0050	34.5115	0.9131
1977	−0.4786	−0.0055	38.9988	0.9580	2002	−0.3399	−0.0051	33.0057	0.9343
1978	−0.3688	−0.0054	33.9461	0.9624	2003	−0.3917	−0.0051	29.6733	0.9119
1980	−0.2926	−0.0053	30.3995	0.9454	2005	−0.3712	−0.0049	34.5907	0.9297
1981	−0.3195	−0.0053	31.5630	0.9572	2006	−0.3172	−0.0053	32.3115	0.9073
1982	−0.3378	−0.0053	32.0980	0.9495	2007	−0.5041	−0.0053	40.3772	0.8963
1983	−0.2558	−0.0052	28.9603	0.9495	2008	−0.5018	−0.0053	40.6206	0.9026
1984	−0.2848	−0.0053	30.3356	0.9484	2009	−0.3835	−0.0051	35.9925	0.9004
1985	−0.2357	−0.0054	27.7359	0.9385	2010	−0.2507	−0.0049	28.8941	0.8953
1986	−0.2469	−0.0054	28.1141	0.9466	2011	−0.5213	−0.0049	40.2703	0.9220
1987	−0.4437	−0.0053	36.1478	0.9477	2012	−0.3468	−0.0049	33.2220	0.8841
1988	−0.3944	−0.0052	34.4490	0.9465	2013	−0.2188	−0.0051	27.5949	0.9057
1989	−0.4365	−0.0050	36.5253	0.9404	2014	−0.3393	−0.0052	32.6235	0.9344
1990	−0.3841	−0.0047	34.1961	0.9075	2015	−0.3847	−0.0052	34.5167	0.9304
1991	−0.1963	−0.0049	26.3644	0.8793	2016	−0.3059	−0.0053	30.7943	0.9098
1992	−0.2019	−0.0051	25.9778	0.9254	2017	−0.3073	−0.0052	30.2121	0.9135
1993	−0.3923	−0.0051	33.7914	0.8973	2018	−0.3663	−0.0052	32.8916	0.9105
1994	−0.2254	−0.0052	26.8465	0.8742	2019	−0.1892	−0.0054	25.5850	0.9376
1995	−0.2914	−0.0053	29.7609	0.9405	2020	−0.2804	−0.0054	28.9743	0.9237
1996	−0.4033	−0.0053	34.5481	0.9553	2021	−0.3977	−0.0053	33.8992	0.9146
1997	−0.4426	−0.0052	36.1949	0.9492	2022	−0.4922	−0.0052	38.0545	0.9254

All parameters significant at 99%

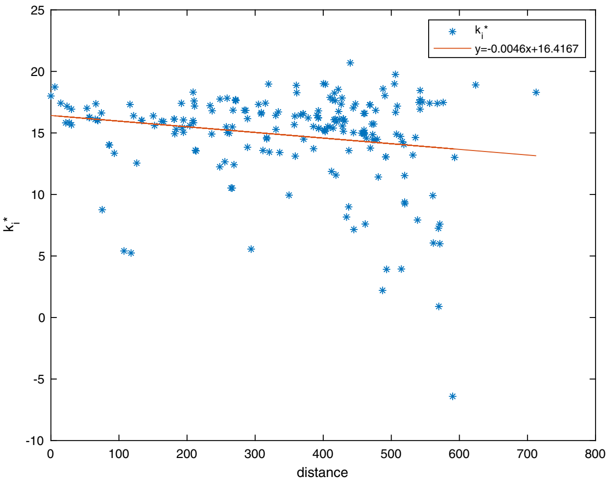

Fig. 4 Linear relation between the distance of the $i-$th station to the REF station and k_i^*, Year 2021

Table 5 Regression analysis between altitude and optimal strike (HDD contracts) for all meteorological stations: 1973-2022

Year	Slope	Intercept	R^2	Year	Slope	Intercept	R^2
1973	−0.0055	16.9737	0.5480	1998	−0.0053	17.6823	0.6352
1974	−0.0055	17.2399	0.6384	1999	−0.0050	17.9945	0.5889
1975	−0.0051	16.6984	0.5836	2000	−0.0052	17.9439	0.5741
1976	−0.0051	16.1404	0.5220	2001	−0.0057	18.4397	0.6295
1977	−0.0054	17.0231	0.5224	2002	−0.0050	18.1899	0.5397
1978	−0.0054	17.3909	0.4971	2003	−0.0056	16.0060	0.6119
1979	−0.0051	17.5190	0.5025	2004	−0.0054	18.1801	0.5808
1980	−0.0052	17.9768	0.5681	2005	−0.0052	18.0410	0.5705
1981	−0.0051	18.1341	0.5303	2006	−0.0054	18.1751	0.5299
1982	−0.0054	17.2855	0.5198	2007	−0.0053	18.4357	0.5813
1983	−0.0050	17.6273	0.5310	2008	−0.0052	19.1692	0.6031
1984	−0.0053	17.7230	0.5672	2009	−0.0055	18.1040	0.5656
1985	−0.0051	17.5899	0.4556	2010	−0.0057	17.5469	0.5263

4.2 Out-of-sample analysis

The validity of the results obtained above is confirmed by the out-of-sample analysis for both HDD and CDD contracts. The data set is divided in proportion 60-40%, that is between training and validation. On the basis of the first 30 years, yearly k_i^* are calculated. Through a multiple linear regression which considers alt_i and lat_i as independent variables, the regression plane is found. Finally, the estimated historical values are used to calculate future values of strikes $\tilde{k}_i^*$.

Before describing the results for the totality of the meteorological stations, we focus on one station and on the CDD contract. Brennero station is relevant because of its distance

Table 5 continued

Year	Slope	Intercept	R^2	Year	Slope	Intercept	R^2
1986	−0.0055	17.7763	0.5079	2011	−0.0053	17.8279	0.5256
1987	−0.0052	17.4773	0.4996	2012	−0.0053	18.1016	0.5333
1988	−0.0053	18.0501	0.6312	2013	−0.0054	18.1834	0.5669
1989	−0.0045	18.2304	0.4556	2014	−0.0056	18.3847	0.6469
1990	−0.0051	17.3703	0.5386	2015	−0.0055	17.9462	0.6702
1991	−0.0050	16.6083	0.4605	2016	−0.0053	17.8570	0.5613
1992	−0.0048	16.5686	0.5403	2017	−0.0053	17.7760	0.5671
1993	−0.0048	16.9465	0.5038	2018	−0.0056	17.8086	0.6203
1994	−0.0055	17.0005	0.5850	2019	−0.0054	18.0432	0.5831
1995	−0.0053	17.8825	0.5916	2020	−0.0053	18.6328	0.6307
1996	−0.0056	17.1797	0.5511	2021	−0.0057	18.0932	0.6037
1997	−0.0054	17.4999	0.5846	2022	−0.0051	17.9528	0.6021

All parameters significant at 99% level

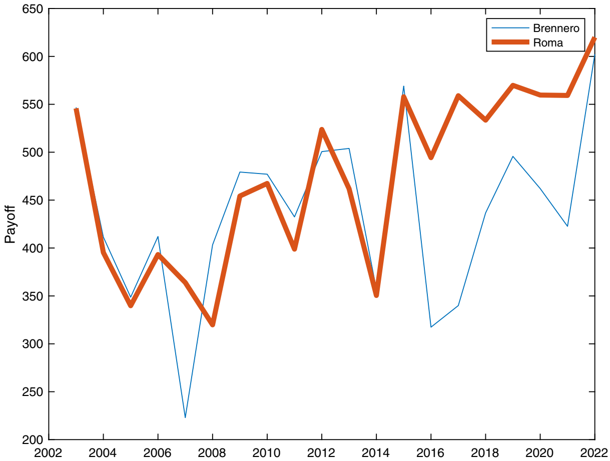

Fig. 5 Comparison between pay-offs of Rome and Brennero station in the out-of-sample period

from Rome, approximately 570 km, and its height above the sea level, 1362 m, representing a location with features very different from those of the reference one.

As shown in Fig. 5, also in the out-of-sample context, the pay-off of the reference derivative is replicated by the WDs with its optimal strike almost perfectly in 12 years out of 20 (from 2003 to 2006, from 2009 to 2015 and in 2022); over 5 years, 2008 and from 2018 to 2021, the approximation error is small, around more or less 10%. Only over 3 years, 2007, 2016 and 2017, the approximation error is about 30%; nevertheless, this case occurs rarely and the magnitude of the deviation from the benchmark is acceptable. This result is remarkable and comparable to what found, for rainfall WDs, in Ritter et al. (2014).

In terms of overall results of this analysis, box-plots in Fig. 6, for CDD contracts, and Fig. 7, for HDD ones, resume the findings for all stations under comparison. In both cases the approximation's effectiveness is evident: for each year from 2003 to 2022, 50% of the

Table 6 Multiple linear regression between altitude, latitude, and optimal strike price-1973-2022-HDD-All parameters significant at 99% level

Year	alt	lat	Intercept	R^2	Year	alt	lat	Intercept	R^2
1973	−1.0262	−0.0044	60.3464	0.9164	1998	−0.8067	−0.0044	51.7760	0.9241
1974	−0.8239	−0.0046	52.0618	0.9171	1999	−0.8100	−0.0041	52.2267	0.8879
1975	−0.8648	−0.0042	53.2465	0.9018	2000	−0.9039	−0.0042	56.1473	0.9119
1976	−0.9834	−0.0040	57.7042	0.9025	2001	−0.9066	−0.0047	56.7549	0.9395
1977	−1.0757	−0.0043	62.4882	0.9213	2002	−0.9343	−0.0040	57.6759	0.9030
1978	−1.1563	−0.0042	66.2629	0.9314	2003	−0.9210	−0.0046	54.9314	0.9322
1979	−1.0698	−0.0040	62.7312	0.9274	2004	−0.9330	−0.0044	57.6107	0.9202
1980	−0.9428	−0.0043	57.8251	0.9226	2005	−0.9258	−0.0043	57.1671	0.9145
1981	−0.9787	−0.0040	59.4974	0.9139	2006	−1.0216	−0.0043	61.3532	0.8992
1982	−1.0729	−0.0042	62.6313	0.9221	2007	−0.8493	−0.0044	54.3308	0.8719
1983	−0.9498	−0.0040	57.7715	0.9046	2008	−0.8154	−0.0044	53.6323	0.8873
1984	−0.9648	−0.0043	58.5009	0.9311	2009	−0.9903	−0.0045	59.9584	0.9210
1985	−1.1671	−0.0038	66.9180	0.9240	2010	−1.1490	−0.0045	66.1080	0.9359
1986	−1.1271	−0.0043	65.4127	0.9260	2011	−1.0662	−0.0042	62.8912	0.9294
1987	−1.1081	−0.0041	64.3122	0.9288	2012	−0.9804	−0.0043	59.5378	0.8884
1988	−0.8351	−0.0045	53.3466	0.9304	2013	−0.9574	−0.0044	58.6464	0.9136
1989	−0.9361	−0.0035	57.7926	0.8388	2014	−0.8322	−0.0047	53.5547	0.9238
1990	−0.9646	−0.0041	58.1368	0.9101	2015	−0.7527	−0.0047	49.7579	0.9159
1991	−1.0892	−0.0039	62.6423	0.8773	2016	−0.9247	−0.0043	56.9394	0.8916
1992	−0.8388	−0.0040	52.0177	0.8525	2017	−0.9152	−0.0044	56.4567	0.8891
1993	−0.9039	−0.0038	55.1479	0.8526	2018	−0.8779	−0.0046	54.9101	0.9183
1994	−0.9241	−0.0045	56.0568	0.9060	2019	−0.9074	−0.0044	56.3920	0.9029
1995	−0.9158	−0.0044	56.5870	0.9259	2020	−0.7684	−0.0045	51.1079	0.8870
1996	−1.0560	−0.0045	61.8095	0.9329	2021	−0.9741	−0.0047	59.2606	0.9419
1997	−0.9480	−0.0044	57.5683	0.9334	2022	−0.8034	−0.0042	51.9062	0.8935

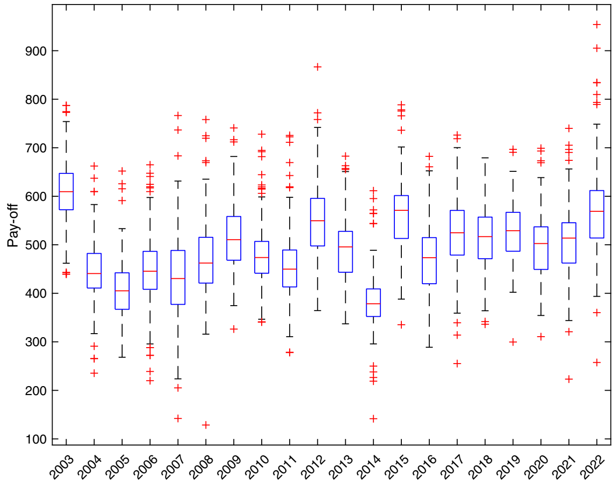

Fig. 6 Box-plot representing, for each year 2003–2022, the out-of-sample distribution of CDD pay-off for all 197 meteorological stations

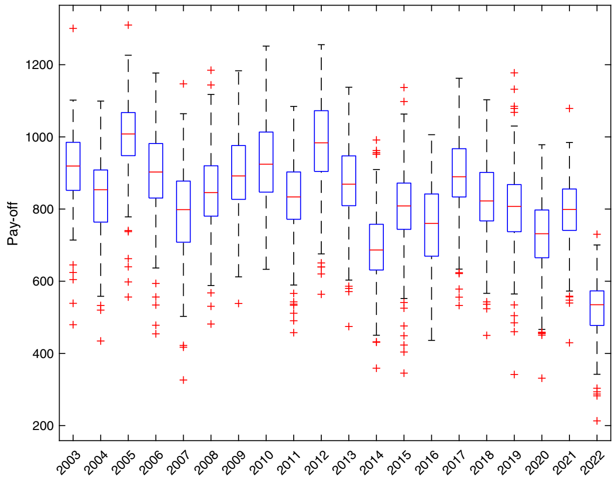

Fig. 7 Box-plot representing, for each year 2003–2022, the out-of-sample distribution of HDD pay-off for all 197 meteorological stations

weather stations display an approximating payoff that deviates from the reference one for less than 10%.

The procedure allows to validate the effective capacity of replication with respect to the reference contract, which, in our analysis, is the only available and quoted contract.

Box plots in Figs. 6 and 7 show the degree of goodness in the out-of-sample replication for the pay-off of the reference contract. The replicated contracts, in the majority of cases, show pay-offs that differ by few percentage points from the reference pay-off, except for some physiological outliers.

5 Conclusions

GBR is the deviation of local meteorological conditions from those of a reference measurement station. This is a strong and intuitive argument against the wide diffusion of WDs as effective risk management tools.

Focusing on the Italian case, this paper shows how, at least for CDD and HDD contracts, the distance between the reference station and the subscriber's location turns out not to be a relevant variable. In particular, Fig. 4 shows an interesting phenomenon, possibly due to the coexistence, with respect to the distance between meteorological stations and the REF one, of two different behaviours. However, further analysis on this circumstance is beyond the scope of this work and left for future research.

Locations' altitudes, instead, reveal to be very useful to adapt traded contracts to local needs. Latitude is also statistically relevant.

In this framework, the low diffusion and use of WDs can be ascribed to a misleading perception of the GBR more than an effective riskiness. Moreover, even though the large level of standardization of WDs contracts could be perceived as a further element that impacts the effectiveness of these products, our methodology bypasses this issue.

A necessary assumption is the existence of a traded reference WD. The optimization procedure makes WDs written on local temperatures as similar as possible to the traded one. Such instruments are then capable of hedging risk in a specified, however large, area. Then, we would expect an increase in the demand of WDs, since their market is more liquid.

It would be interesting to investigate if the impact of GBR is reduced also in presence of a different underlying, such as wind or rainfall. This is left for further research.

Acknowledgements The authors of this article gratefully thank Prof. Luigi Mariani and Prof. Franco Zavatti. Their excellent meteorological data management is here warmly acknowledged very useful suggestions by an anonymous referee are fully acknowledged.

Author contributions All authors contributed equally to this work.

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement. This study received no funding.

Declarations

Conflict of interest all Authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Abdi, M. J., Raffar, N., Zulkafli, Z., et al. (2022). Index-based insurance and hydroclimatic risk management in agriculture: A systematic review of index selection and yield-index modelling methods. *International Journal of Disaster Risk Reduction, 67*(102), 653. https://doi.org/10.1016/j.ijdrr.2021.102653

Alexandridis, A., & Zapranis, A. D. (2012). *Weather derivatives: Modeling and pricing weather-related risk*. Springer.

Benth, F. E., & Benth, J. V. (2007). The volatility of temperature and pricing of weather derivatives. *Quantitative Finance, 7*(5), 553–561. https://doi.org/10.1080/14697680601155334

Bucheli, J., Dalhaus, T., & Finger, R. (2021). The optimal drought index for designing weather index insurance. *European Review of Agricultural Economics, 48*(3), 573–597. https://doi.org/10.1093/erae/jbaa014

Castellano, R., Cerqueti, R., & Rotundo, G. (2020). Exploring the financial risk of a temperature index: A fractional integrated approach. *Annals of Operations Research, 284*(1), 225–242. https://doi.org/10.1007/s10479-018-3063-0

East, M. (2005). Issues of geographical basis risk in weather derivatives for Australian wheat farmers. Tech report

Figini, S., & Uberti, P. (2010). Model assessment for predictive classification models. *Communications in Statistics-Theory and Methods, 39*(18), 3238–3244. https://doi.org/10.1080/03610920903243751

Franzoni, S., & Pelizzari, C. (2021). Rainfall option impact on profits of the hospitality industry through scenario correlation and copulas. *Annals of Operations Research, 299*(1), 939–962. https://doi.org/10.1007/s10479-019-03442-5

Gyamerah, S. A., Ngare, P., & Ikpe, D. (2019). Mitigating geographical basis risk of weather derivatives using spatial-temporal regime-switching temperature model. *AIMS Mathematics, 4*(4), 1274–1290. https://doi.org/10.3934/math.2019.4.1274

Hellmuth, M. E., Osgood, D. E., Hess, U., et al. (2009). *Index insurance and climate risk: Prospects for development and disaster management*. Technical report: International Research Institute for Climate and Society (IRI).

Jewson, S., & Brix, A. (2005). *Weather derivative valuation: the meteorological, statistical, financial and mathematical foundations*. Cambridge: Cambridge University Press.

Musshoff, O., Odening, M., & Xu, W. (2011). Management of climate risks in agriculture-will weather derivatives permeate? *Applied Economics, 43*(9), 1067–1077. https://doi.org/10.1080/00036840802600210

Ritter, M., Musshoff, O., & Odening, M. (2014). Minimizing geographical basis risk of weather derivatives using a multi-site rainfall model. *Computational Economics, 44*(1), 67–86. https://doi.org/10.1007/s10614-013-9410-y

Stefani, S., Moretto, E., Parravicini, M., et al. (2018). Managing adverse temperature conditions through hybrid financial instruments. *Journal of Energy Markets, 11*(3), 25–41. https://doi.org/10.21314/JEM.2018.182

Stefani, S., Kutrolli, G., Moretto, E., et al. (2020). Managing meteorological risk through expected shortfall. *Risks, 8*(4), 118. https://doi.org/10.3390/risks8040118

Taib, C. M. I. C., & Benth, F. E. (2012). Pricing of temperature index insurance. *Review of Development Finance, 2*(1), 22–31. https://doi.org/10.1016/j.rdf.2012.01.004