Two-stage stochastic optimization for industrial heat systems under crisis-driven price uncertainty

Karim DARWICH [a], François ROUSSET [b], Benoît CHEVALIER-ROIGNANT [c], Marc CLAUSSE [d]

[a] INSA Lyon, CNRS, CETHIL, UMR 5008, Lyon, France, karim.darwich@insa-lyon.fr, CA

[b] INSA Lyon, CNRS, CETHIL, UMR 5008, Lyon, France, francois.rousset@insa-lyon.fr

[c] emlyon business school, Lyon, France, benoit.chevalier-roignant@outlook.com

[d] INSA Lyon, CNRS, CETHIL, UMR 5008, Lyon, France, marc.clausse@insa-lyon.fr

Abstract:

Decarbonizing industrial heat requires long-term capacity investments under uncertainty, where energy price dynamics are often structurally correlated. As highlighted by the 2021–2022 European energy shock and ongoing geopolitical tensions, these investment decisions must be made under uncertainty in both price levels and their co-movement. Our approach contrasts with most existing planning studies that rely on deterministic optimization (which ignores price uncertainty entirely) or scenario-based Monte Carlo approaches that assume independent commodity dynamics. The limitation of these studies lies not in the use of scenarios per se, but in the assumption of independent or Gaussian price dynamics, which assigns low probability to simultaneous gas–electricity crises. This paper proposes a two-stage stochastic optimization framework for the optimal sizing and technology selection of industrial heat supply systems that combines three methodological elements not widely used in the industrial heat planning literature: (i) jump-driven price dynamics (Sum of Ornstein–Uhlenbeck processes) calibrated to historical gas and electricity data; (ii) a 4-dimensional t-copula scenario generation procedure that explicitly represents heavy-tailed joint dependence among gas prices, electricity prices, solar irradiance, and ambient temperature; and (iii) a mean–CVaR weighted two-stage stochastic linear program (LP) to determine the minimal expected cost and maximal tail-risk protection.

The central question addressed is: given uncertainty about future energy prices, including their tendency to co-move during crises, what industrial heat technology portfolio minimizes long-run cost while remaining robust to severe outcomes? We compare the deterministic expected-value (EV) baseline model, which optimizes at mean prices, with two-stage stochastic models for (i) risk-neutral and (ii) risk-averse models. The results demonstrate that deterministic planning underestimates system risk and leads to different investment decisions compared to the stochastic formulation.

Keywords:

Decarbonization; Electrification; Industrial Heat Systems; Jump-diffusion Processes; Stochastic Programming; Uncertainty.

1. Introduction

Industrial process heat accounts for approximately 42% of global final energy consumption and over 20% of global CO_2 emissions, yet it remains one of the least-addressed segments of the energy transition [1]. A significant share of this consumption, particularly processes requiring continuous thermal supply at temperatures between 80°C and 150°C (characteristic of food sterilization, pharmaceutical manufacturing, and paper drying), remains a section with limited electrification and exposed to fossil fuel price uncertainty. The European Commission has identified industrial heat decarbonization as a strategic priority under the

REPowerEU plan, with heat pumps, electric boilers, and solar thermal collectors as primary substitution technologies [2].

Decarbonizing these processes requires long-term capital investments that will be operated over 20-year horizons under uncertain future energy prices. The design challenge is fundamentally economic: the optimal choice of technology portfolio depends on how accurately future energy prices can be predicted. Once committed, these large-scale investments are hardly reversible [3], so that an operator who mis-invests in dispatchable backup cannot readily adjust its production portfolio should energy prices shift dramatically after the investment is committed, an operator who invested wisely in fuel-flexible capacity may gain a decisive competitive advantage, whereas one locked into a single-fuel strategy faces unabsorbable cost spikes. The 2021–2022 European energy crisis illustrated this risk directly: gas and electricity prices co-moved sharply, exposing operators who had relied on single-fuel strategies to simultaneous input cost spikes they could not absorb [4] a pattern that continues to be observed amid ongoing geopolitical tensions affecting global energy markets. McKinsey (2026) projects a further 60%–170% increase in European electricity price volatility by 2030 as renewable penetration rises, making the problem more critical over the planning horizon of investments made today [5].

The literature on energy system planning under uncertainty has evolved substantially over the past decade, producing sophisticated frameworks for district heating networks [6], distributed energy systems, and power systems [7]. Despite this progress, the industrial heat sub-sector lags behind in its treatment of uncertainty. Most existing studies on industrial heat supply design rely on one of three approaches: deterministic techno-economic optimization at assumed price points ([1], [8]); simple Monte Carlo sampling with independent commodity price paths [9]; or scenario-tree formulations that capture temporal price variation but not cross-commodity tail dependence [7]. Each approach, however, has limitations. In particular, deterministic models [1], [8] commit the expected-value fallacy [10]: they optimize for average prices and might underinvest (by Jensen's inequality, optimizing at the mean of a convex recourse cost function yields a lower-bound estimate of the true expected cost, leading to a systematically smaller backup capacity than is optimal under uncertainty) in backup capacity that only pays off in rare but severe conditions [10]. Independent stochastic models (e.g., [9]) correctly reproduce each commodity's marginal distribution but underestimate the probabilities of episodes during which gas and electricity are both expensive. Scenario-tree formulations (e.g., [7]) can capture temporal structure but typically rely on Gaussian or independent copula (a function that links individual price distributions into a joint one) that fail to reproduce the tendency of energy prices to spike together in extreme market conditions [11], [12].

Our present work applies a two-stage stochastic optimization framework for optimizing the sizing and technology selection of industrial heat supply systems. Our research extends previous works by combining three methodological elements rarely used together in industrial heat planning: (i) a multidimensional jump-diffusion process for modeling the price dynamics which is calibrated to historical gas and electricity price data; (ii) a *t*-copula scenario generation procedure, where the Student-t copula explicitly represents heavy-tailed joint dependence between energy carriers, ensuring that joint price extremes are reasonably represented; and (iii) a "mean-CVaR weighted objective," in which Conditional Value-at-Risk (CVaR) measures the expected cost in the worst fraction of scenarios, allowing the planner to trade off average performance against tail-risk exposure. In summary, this study contributes to the field in the following ways:

- Jump-diffusion price processes (Sum of Ornstein–Uhlenbeck processes for electricity, Merton jump-diffusion for gas) were calibrated to historical European data, accurately reproducing both marginal distributions and extreme tail behavior.
- To model the structural gas-electricity co-movement observed during the 2021–2022 crisis, a 4-dimensional *t*-copula scenario generation framework is proposed.

The remainder of this paper is structured as follows: Section 2 describes the facility and its decision problem. Section 3 presents the integrated modeling framework including price process calibration, copula scenario generation, and the two-stage stochastic program. Section 4 presents validation results. Section 5 presents the results and comparative analysis. Finally, Section 6 summarizes managerial insights and suggests future research directions.

2. Energy system model

2.1. Modelling overview

The model determines optimal long-term capacity allocation across a portfolio of industrial heat supply technologies for a single continuous-process facility. The objective is to minimize cost over a 20-year horizon, combining annualized capital expenditure (CAPEX), fixed Operations and Maintenance (O&M) and variable operating costs (OPEX). The model represents a single thermal node delivering heat at 100°C. At the hourly resolution, the model operates at hourly resolution within 12 representative weeks selected via peak-preserving *k*-medoids clustering to preserve annual distributions of prices, solar irradiance, and temperature.

2.2. System topology

Figure 1 illustrates the flow: electricity grid, natural gas network, and solar irradiance as inputs; heat pump (HP), gas boiler (GB), electric boiler (EB), solar thermal (ST), and thermal energy storage (TES) as conversion/storage technologies. The heat balance at each hour is given by Eq. (1):

$$n_{HP}(t) + n_{GB}(t) + n_{ST}(t) + n_{EB}(t) + tes_{dis}(t) - tes_{ch}(t) + load_{shed}(t) = demand(t) = 1{,}500kW \quad (1)$$

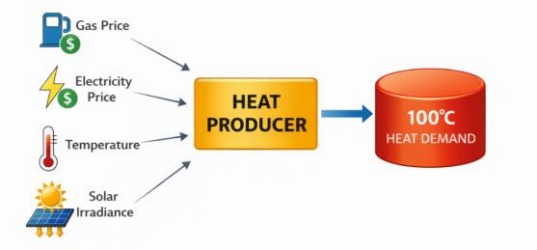

Figure. 1. *Simple schematic of system topology of the industrial heat supply model.*

The facility operates continuously at a fixed thermal demand of $\dot{Q} = 1{,}500\ kW$ at a delivery temperature of $T_{sink} = 100°C$. Load shedding carries an economic penalty of $€10/kWh$, consistent with the upper range of production-disruption cost estimates for European manufacturing, reflecting the direct costs of lost production and contractual penalties [13].

2.2.1. Energy carriers

The model considers two stochastic energy carriers whose prices vary across scenarios as described in Section 3. Electricity is procured from the grid at time-varying hourly spot prices; natural gas is purchased at monthly step prices representing seasonal contract structures. The grid electricity mix carries an emission factor of $0.10\ kgCO_2/kWh$, reflecting an intermediate value between the French average grid intensity ($0.052\ kgCO_2/kWh_{el}$ [14]) and the emission rate of gas-fired marginal generation [15]. Natural gas combustion emits $0.202\ kgCO_2/kWh_{th}$ [16]. A carbon price of $€65/\ tCO_2$ is applied as a fixed charge, consistent with EU ETS observed spot prices over 2022–2024 [17], and enters the objective function through the per-technology CO_2 emission rate.

2.2.2. Technology portfolio and technical constraints

The facility may invest in five technologies jointly; key technical parameters are summarized in Table 1.

Table. 1. *Techno-economic parameters of the candidate technologies. [7], [18], [19]*

Technology	CAPEX (€/kW)	Fixed O&M (€/kW·yr)	Lifetime (yr)	Key parameter	Value
HP	1,000	3.0	20	η_2 (2nd law eff.)	0.50
ST (solar thermal)	600	1.5	25	$\frac{\eta_{opt}}{U_L}$ (W/m²K)	0.75/2.0
GB (gas boiler)	110	3.0	25	Combustion efficiency	0.92
EB (electric boiler)	120	0.5	20	Electrical efficiency	0.98
TES	30 (€/kWh)	0.5	25	Round-trip efficiency	0.9025 (0.95^2)

The heat pump (HP) converts electricity to heat via a vapor-compression cycle. Its coefficient of performance (COP) is computed following [18] as a Carnot COP adjusted by a second-law efficiency $\eta_2 = 0.50$ and given by Eq. (2)

$$COP(t) = \eta_2 \cdot \frac{T_{sink}}{T_{sink} - T_{source(t)}}, \quad (2)$$

where temperatures are in K units and $T_{source}(t)$ is the hourly ambient temperature, drawn from the Meteostat hourly dataset for Lyon, France (2015-2024). At the reference mid-latitude European climate used here, the mean annual COP is approximately 2.07, with a seasonal range of 1.70 and 2.50.

The gas boiler (GB) converts natural gas to heat at a fixed thermal efficiency of $\eta_{GB} = 0.92$. It provides fully dispatchable backup with no weather dependence.

The electric boiler (EB) converts electricity to heat at $\eta_{EB} = 0.98$, providing fast-response electric backup primarily activated when gas prices are elevated but electricity remains affordable.

The solar thermal collector (ST) produces heat according to the Hottel–Whillier Eq. (3) [20] (for the sake of simplicity, only direct heat production with ST has been considered).

$$q_{ST}(t) = \eta_{opt} \cdot G(t) - U_L \cdot \left(T_{sink} - T_{amb}(t)\right), \quad (3)$$

where G(t) is the plane-of-array global irradiance (W/m²), $\eta_{opt} = 0.75$ is the optical efficiency, and $U_L = 2.0 \frac{W}{m^2 \cdot K}$ is the heat loss coefficient.

The thermal energy storage (TES) system is modelled as a sensible-heat water tank with separate power capacity C_{TES_P} (kW) and energy capacity C_{TES_E} (kWh). The state-of-charge evolves as given in Eq. (4):

$$SOC(t+1) = SOC(t) + \eta_{ch} \cdot tes_{ch}(t) - \frac{tes_{dis}(t)}{\eta_{dis}}, \quad (4)$$

with $\eta_{ch} = \eta_{dis} = 0.95$ (round-trip efficiency 0.9025) and $SOC \in [0, C_{TES_E}]$.

All technologies are evaluated over a uniform 20-year planning horizon at a discount rate of 8%, consistent with the weighted average cost of capital (WACC) for industrial energy investments in European techno-economic assessments [19].

3. Uncertainty modelling and optimization framework

The uncertainty modelling framework is designed to answer one question: given uncertainty about future energy prices (including their tendency to co-move in crises), what industrial heat technology portfolio minimizes long-run cost while remaining robust to severe outcomes? **Figure 2** illustrates the end-to-end modeling framework.

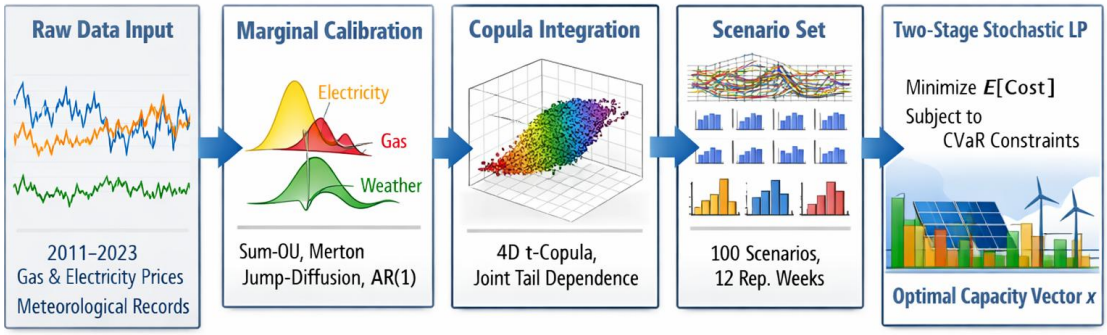

Figure. 2. *End-to-end modelling framework.*

3.1. Optimization framework

3.1.1. Problem structure

The investment problem is formulated as a two-stage stochastic linear program (LP). First-stage decisions define installed capacities $x \in \mathbb{R}$, taken prior to uncertainty realization. Second-stage decisions correspond to hourly dispatch for each scenario $s = 1, \ldots, N$, ensuring demand satisfaction at minimum operating cost under physical constraints.

3.1.2. Objective function

The risk-adjusted objective minimizes annualized investment cost, fixed O&M, and a weighted combination of expected operating cost and Conditional Value-at-Risk (CVaR) given by the weight $\beta \in [0,1]$ controls the trade-off: $\beta = 0$ is risk-neutral; $\beta \to 1$ increasingly penalizes tail outcomes. The random vector ξ = (ξ_{elec}, ξ_{gas}, ξ_{temp}, ξ_{solar}) represents scenario-specific price and weather realizations; CVaR[ξ] is the expected operating cost in the worst (1−α) fraction of scenarios.

$$min\{CRF \cdot CAPEX(x) + Fixed_{OM}(x) + (1-\beta) \cdot \Sigma_s \pi_s \cdot OPEX_s(x, \xi_s) + \beta \cdot CVaR_\alpha[OPEX(x, \xi)]\}, \quad (5)$$

where β ∈ [0,1] is the risk-aversion weight, α = 0.90 is the CVaR confidence level, $\xi_s = \left(elec_{price_s}, gas_{price_s}, temperature_s, solar_s\right)$ is the vector of stochastic parameters for scenario s, and $\pi_s = \frac{1}{N}$ are equal scenario probabilities.

Why CVaR rather than expected cost alone? Minimizing expected cost (β = 0) is risk-neutral: it weights a €1 cost in a crisis year identically to a €1 cost in a normal year. CVaR at confidence level α is defined as the expected cost conditional on being in the worst (1−α) fraction of scenarios: $CVaR_{90}$ is the average cost across the 10 worst scenarios out of 100. It is a coherent risk measure [21], [22] satisfying monotonicity, sub-additivity, homogeneity, and translation invariance, properties that expected cost and Value-at-Risk (VaR) do not jointly possess. Unlike VaR, which only reports a threshold, CVaR explicitly penalizes the severity of outcomes

beyond that threshold. The weight β allows the operator to interpolate continuously between risk-neutral (β = 0) and tail-risk-minimizing (β → 1) objectives; Section 5 characterizes the cost of protection across the mean-CVaR efficiency frontier (the set of Pareto-optimal portfolios that cannot improve CVaR without increasing expected cost).

CVaR is linearized using the [22] auxiliary-variable reformulation using the following Eq. (6):

$$CVaR_\alpha = \zeta + \left[\frac{1}{1-\alpha}\right] \cdot \Sigma_s \pi_s \cdot \eta_s, with \eta_s \geq OPEX_s - \zeta, \eta_s \geq 0, \zeta \in R \quad (6)$$

where ζ is the auxiliary decision variable representing the α-quantile of the operating cost distribution, optimized jointly with x, and η_s is the excess of scenario s above ζ [22].

3.1.3. Per-scenario operating cost and emissions

The per-scenario annual operating cost $OPEX_s$ is the week-weighted annual sum of fuel, carbon, and load-shedding costs using Eq. (7):

$$OPEX_s(x, \xi_s) = \Sigma_{\{k=1\}}^{\{K\}} w_{\{s,k\}} \cdot \Sigma_{\{t=1\}}^{\{168\}} \left[p_{gas}(s,k) \cdot \frac{gen_{GB}(s,k,t)}{\eta_{GB}} + p_{elec}(s,k,t) \cdot \left(\frac{gen_{HP}(s,k,t)}{COP(s,k,t)} + \frac{gen_{EB}(s,k,t)}{\eta_{EB}} \right) + c_{carbon} \cdot e_s(k,t) + VOLL \cdot load_{shed}(s,k,t) \right], \quad (7)$$

where $w_{\{s,k\}}$ is the integer cluster weight of representative week *k* in scenario s ($\Sigma_k w_{\{s,k\}} = 52$ for all s). Because temporal clustering is applied independently per scenario (Section 3.3), the representative weeks and their weights differ across scenarios, making the double index necessary. The hourly CO_2 emissions rate is given by

$$e_s(k,t) = f_{gas} \cdot \frac{gen_{GB}(s,k,t)}{\eta_{GB}} + f_{grid} \cdot \left[\frac{gen_{HP}(s,k,t)}{COP(s,k,t)} + \frac{gen_{EB}(s,k,t)}{\eta_{EB}} \right] \quad (8)$$

where $f_{gas} = 0.202$ kgCO_2/kWh (IPCC, 2006) and $f_{grid} = 0.10$ kgCO_2/kWh (Section 2).

3.1.4. Key performance indicators

Results are reported against four KPIs:

1. The Levelized Cost of Heat (**LCOH**) is the ratio of total annualized system cost to annual heat delivered (13,104 MWh/yr), expressed in €/MWh_{th}; it is the primary cost-efficiency metric for comparing planning approaches.
2. The Value of the Stochastic Solution (**VSS** = EEV − RP) quantifies the annual economic benefit of stochastic over deterministic planning [10], where EEV is the cost incurred by the deterministic solution evaluated on the full scenario set and RP is the recourse problem optimum.
3. The Expected Value of Perfect Information (**EVPI** = RP − WS), where WS is the wait-and-see solution, provides an upper bound on the value of resolving uncertainty entirely [10]
4. The **$CVaR_{90}$** of operating cost directly measures tail-cost exposure and serves as the primary risk metric for comparing copula specifications and risk-aversion levels.

3.2. Capturing electricity and gas price dynamics

3.2.1. The physical origin of price jumps in energy markets

Electricity prices exhibit spikes due to non-storability and inelastic demand, while gas prices experience jumps from geopolitical supply shocks (Figure 3).

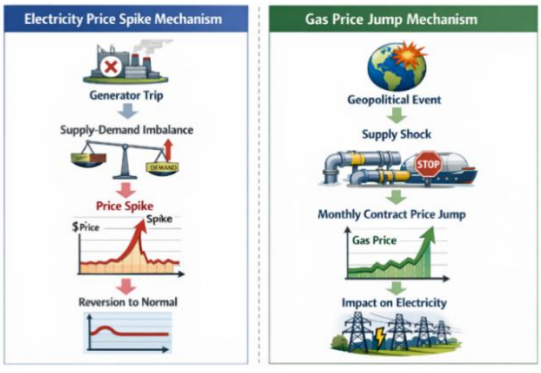

Figure. 3. Physical origins of energy price jumps. Left: electricity price spikes arise from the non-storability of power and inelastic short-run demand, producing fast-reverting spikes captured by the Sum-OU process. Right: natural gas price jumps reflect discrete geopolitical supply shocks and contractual step-changes, modelled as a Merton compound Poisson process

3.2.2. Electricity prices: sum of two Ornstein–Uhlenbeck processes

An Ornstein–Uhlenbeck (OU) process is a mean-reverting stochastic process in which prices are continuously pulled back toward a long-run equilibrium level after any deviation, capturing the physics of electricity markets where scarcity spikes are followed by demand destruction and supply re-entry. To reproduce both the slow weekly base dynamics and the fast intra-day spike behavior of electricity prices, two distinct mean-reverting components are required. A single OU process cannot simultaneously match both time scales: calibrating it to weekly autocorrelation produces a process too smooth to generate intra-day spikes, while calibrating it to hourly spikes produces a process reverting too quickly to capture weekly persistence. The Sum-OU specification, following [23] and [24], resolves this by superimposing two independent OU processes with different mean-reversion speeds using the relationship

$$X(t) = f(t) + Y_1(t) + Y_2(t), \quad (9)$$

where f(t) is a deterministic Fourier seasonality component capturing annual, semi-annual, weekly, and diurnal harmonics; $Y_1(t)$ is a slow-mean-reverting base process with mean-reversion speed θ_1 = 0.00874 h^{-1} (characteristic time λ_1 half-life 79 h, characteristic time 114 h), capturing multi-day price trends; and $Y_2(t)$ is a fast-mean-reverting spike process with θ_2 = 0.1382 h^{-1} (λ_2 half-life 5.0 h), capturing intra-day spikes that dissipate within a few hours of the supply scarcity event.

3.2.3. Gas prices: Merton jump-diffusion

Gas prices can be modelled as a jump-diffusion, a process that combines a continuous random walk ("diffusion") and sudden discrete jumps representing market shocks. Specifically, we consider a geometric Brownian motion augmented with a compound Poisson process (which models events that arrive randomly but at a statistically stable average rate, such as a gas contract renegotiation that occurs roughly once per month) representing discrete contractual and geopolitical shocks using Eq. (10) [25]:

$$dP_{gas}(t) = \mu P_{gas}(t)dt + \sigma P_{gas}(t).dW(t) + J(t).dN(t) \quad (10)$$

where $W(t)$ is a standard Brownian motion, $N(t)$ is a Poisson process with jump intensity λ, and $J(t)$ is the random jump amplitude. The jump amplitude J(t) follows a log-normal distribution $J \backsim \varepsilon \cdot exp\left(N\left(\mu_J, \sigma_J^2\right)\right)$, where μ_J and σ_J are calibrated from historical gas price series. The compound Poisson process N(t) has jump intensity λ (mean number of jumps per unit time). The Merton specification is appropriate here because gas prices exhibit two clearly separable modes of variation: a slow diffusion component reflecting gradual supply-demand rebalancing, Kou (2002) uses a double-exponential distribution for jump amplitudes, which could be considered as an alternative [26]. and a discrete jump component reflecting the step-change structure of monthly contract renegotiations. The model is calibrated separately from electricity, with jump intensity and amplitude estimated from historical monthly gas price series over 2011–2023.

3.3. Modelling joint price crises: from independent marginals to correlated tail events

3.3.1. Why marginal distributions alone are not enough

Accurate representation of marginal price distributions is necessary but not sufficient for investment planning under multi-energy uncertainty. The relevant quantity for system design is the frequency and severity of joint extreme events, particularly scenarios in which gas and electricity prices are simultaneously elevated. These joint tail events determine the economic value of fuel-flexible technologies.

Under adverse conditions, for instance, electricity ≈ €270/MWhel, gas ≈ €75/MWhth, the marginal cost of heat from the heat pump rises to ≈ €130/MWhth ([computed as: HP marginal $cost = \frac{p_{elec}}{COP} \approx \frac{270}{2.07} \approx 130€/MWh_{th}$; GB marginal $cost = \frac{p_{gas}}{\eta_{GB}} \approx \frac{87}{0.92} \approx €95/MWh_{th}$ at crisis prices]), exceeding the gas boiler ≈ €95/MWh_{th}, thereby reversing the merit order. Models assuming independence typically generate these extremes in separate scenarios and therefore fail to capture this regime shift, leading to structurally biased investment decisions. The dependence structure is thus a determinant of the optimal portfolio, not a secondary modelling choice. A deeper reason to model gas and electricity jointly is the endogeneity of their price dynamics: both are driven by shared macroeconomic fundamentals, global economic activity, geopolitical tensions affecting fossil fuel supply chains, and carbon pricing, even if the specific shock mechanisms differ (geopolitical disruptions, contract renegotiations for gas; demand inelasticity and renewable intermittency for electricity). Treating them as independent processes ignores this common latent driver, underestimating the probability of crisis episodes in which both inputs are simultaneously expensive. The copula approach adopted here specifically addresses this by giving extra statistical weight to joint extreme events, reflecting the empirical observation that severe gas and electricity price spikes tend to co-occur rather than alternate.

3.3.2. Copula theory: separating margins from dependence

A copula is a function, that couples *d* marginal cumulative distribution functions (CDF) into a joint distribution, while leaving each marginal unchanged. By Sklar's theorem [27], any joint distribution $F(x_1, \ldots, x_d)$ can be written as $F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d))$, where F_i are the marginal CDFs and C is unique on the support

of continuous marginals. This separation is precisely what is needed here: the marginal distributions of gas and electricity prices are calibrated independently in Section 3.2, and the copula is then used solely to impose their joint dependence structure, specifically, how the rank of a gas price year relates to the rank of an electricity price year, independently of the shape of either marginal. The choice of copula family is determined by the property of upper tail dependence. Let U_1 and U_2 be the probability-integral transforms of the two-price series, so that each is uniformly distributed on [0,1]. Upper tail dependence is then defined by Eq. (11)

$$\lambda_U = lim_{\{u \to 1\}} P(U_1 > u \mid U_2 > u), \quad (11)$$

λ_U measures the limiting probability that one variable is extreme given that the other is extreme, as events become increasingly severe.

For a Gaussian copula, $\lambda_U = 0$ whenever the two variables are not perfectly correlated, meaning joint extremes vanish asymptotically: as $u \to 1$ in Eq. (11), the conditional probability that one price is extreme given the other is extreme goes to zero. The Gaussian copula therefore underweights the most damaging scenarios, those where gas and electricity prices are simultaneously far above their means.

In contrast, a *t*-copula with finite degrees of freedom ν exhibits strictly positive upper tail dependence. Denoting by $\rho \in [-1,1]$ the Spearman rank correlation between the two-price series, the closed-form expression is given by Eq. (12)

$$\lambda_U = 2t_{\{\nu+1\}}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right) \quad (12)$$

where $t_{\{\nu+1\}}$ is the CDF of the Student-*t* distribution with ν+1 degrees of freedom [28]. The empirical Spearman rank correlation ρ = 0.964 is computed from day-ahead electricity prices and natural gas spot prices over 2015–2024, and setting ν=4 (a commonly used value in the energy finance literature [11]), we obtain $\lambda_U \approx 0.774$, indicating strong asymptotic co-occurrence of extreme events.

In practice, this tail mass translates directly into the scenario fan. A striking visual difference between the two panels of Figure 4 is the diagonal alignment of scenarios in the t-copula model: because the target dependence structure (ρ = 0.964) induces strong rank co-movement, although the realized scenario correlation is lower due to discrete year mapping between gas and electricity, scenarios are constrained to move together in rank space, cheap gas years coincide with cheap electricity years and expensive years co-occur, reproducing the historical co-movement observed in the 2015–2024 calibration data. The independent model, by contrast, permits combinations such as a scenario where gas is at its 90th percentile while electricity is simultaneously near its 20th percentile, a joint configuration that has no counterpart in the observed record. The joint tail cluster visible in the upper-right corner of Figure 4(b) is simply the extreme end of this diagonal band. The crisis region is defined as annual mean gas price exceeding $60\ EUR/MWh_{th}$ and annual mean electricity price exceeding $200\ EUR/MWh_{el}$, thresholds calibrated to the lower bound of the 2021–2022 energy crisis regime, the only years in the 2015–2024 dataset to breach both bounds simultaneously. These values are consistent with documented market outcomes: ACER (2025) reports an EU average day-ahead electricity price of 227 EUR/MWh in 2022 [29], while TTF annual mean gas prices exceeded 100 EUR/MWh in 2022, against a pre-crisis baseline of 20–25 EUR/MWh [30]. The 60 EUR/MWh gas threshold corresponds to the level at which crisis conditions were already recognized by market participants in autumn 2021, before prices continued their ascent toward the intra-crisis peak [30]. Under these definitions, the t-copula places 20 scenarios in the joint crisis region compared to 12 under the independent model, increasing the frequency of joint crisis. This increase propagates to the system cost distribution, raising the CVaR minimized in the second-stage problem and driving the recourse investment in gas-boiler backup capacity. The copula choice therefore impacts both scenario realism and optimal design.

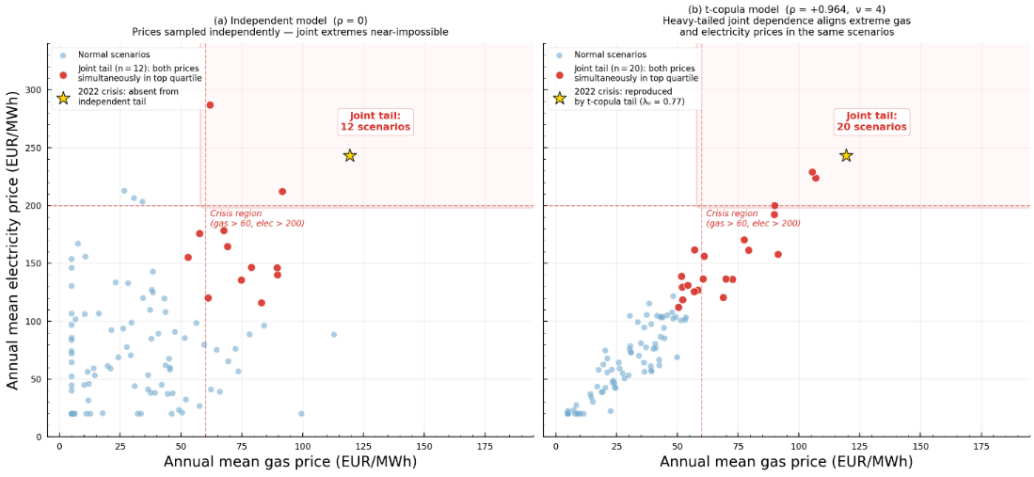

Figure. 4. *Scenario fan: Independent model vs t-copula.*

3.3.3. Scenario generation

A 4-dimensional *t*-copula is fitted jointly to (i) gas price (Eq. (10)), (ii) electricity price (Eq. (9)), (ii) ambient temperature, and (iv) solar irradiance, capturing not only the gas-electricity co-movement but also the weather correlations relevant to the HP COP and solar thermal output. For each of multiple scenarios, a draw $(u_{gas}, u_{elec}, u_{temp}, u_{solar})$ is taken from the fitted copula and mapped to historical price years by rank matching: u_{gas} selects the historical gas price year whose annual rank most closely matches the drawn uniform, and similarly for u_{elec}. This hybrid approach preserves empirical hourly price profiles as marginal inputs, avoiding any parametric assumption on intra-year price shapes, while imposing the calibrated joint tail dependence structure at the annual selection stage.

3.3.4. Latin hypercube sampling

With N = 100 scenarios, Latin Hypercube Sampling (LHS) stratifies each marginal into N equal-probability intervals, guaranteeing coverage of the tails that drive investment decisions [31]. For fixed N, LHS reduces estimator variance relative to Monte Carlo (MC) by eliminating clustering; Section 4 confirms LHS at N = 100 achieves better correlation approximation than MC at N = 200, providing reliable joint crisis scenarios at half the computational cost.

4. Scenario validation

Scenario quality is evaluated across four complementary dimensions, consistent with stochastic programming practice [32]: (i) marginal distribution accuracy, (ii) temporal structure, (iii) cross-variable dependence, and (iv) tail behavior. All statistics are computed from N = 100 scenarios unless otherwise stated. **Table 2** compares simulated and historical marginal statistics.

Table. 2. *Layer 1 validation: marginal distribution statistics for electricity and natural gas (N = 100 scenarios). |Δ| = absolute percentage deviation from historical pool.*

Metric	Historical	Simulated	|Δ|
Electricity (11 years, 2015–2025)			
Mean (€/MWh)	76.64	76.35	0.4%
Standard deviation (€/MWh)	88.46	87.38	1.2%
CVaR (90%) (€/MWh)	292.37	290.80	0.5%
Natural gas (2011–2023 historical pool; 2024 calibration baseline)			
Mean (€/MWh)	30.65	34.66	13.1%
Standard deviation (€/MWh)	27.38	30.85	12.7%
CVaR (90%) (€/MWh)	103.54	116.01	12.0%

Electricity matches historical metrics within 1.2% accuracy ($CVaR_{90}$ gap 0.5%). The gaps between the gas mean (13.1%), and $CVaR_{90}$ (12.0%) are by design: calibration uses a 2024 forward-looking baseline embedding heightened geopolitical risk, rather than the longer pre-crisis pool; VaR_{99} matches exactly (0.0%). The higher mean and CVaR for gas reflect the use of a 2024 forward-looking calibration capturing elevated market volatility. Results should therefore be interpreted relative to this stress-calibrated baseline (This reflects the impact of stress-period calibration on marginal distributions). The Spearman correlation between gas and electricity in the generated scenarios is 0.773 (target 0.964), a structural consequence of mapping continuous copula draws to a discrete pool of historical years; this yields conservative lower bounds on tail exposure.

Scenario set stability is assessed following the independent-replication procedure of [25]. For each sample size N, M independent solves are performed, each drawing a fresh set of N scenarios under a different random seed. Two diagnostics are reported. The coefficient of variation of the recourse problem objective across replications, CV (obj) = σ/μ, measures how sensitive the optimal cost is to the specific scenario draw, a CV below 2% is the accepted convergence threshold in stochastic programming practice. The out-of-sample gap, $OOS_{gap} = \frac{z_{OOS} - z_{IS}}{z_{IS}}$, measures the percentage by which the in-sample optimal cost underestimates the true expected cost when the first-stage decisions are evaluated on a fresh, independent scenario set; a negative OOS gap indicates in-sample optimism, i.e. the model is slightly over-fitted to its training scenarios. **Table 3** reports both metrics across sample sizes.

Table. 3. *Independent-replication stability test.*

N	M	LCOH μ ± σ (€/MWh)	CV (obj)%	OOS gap%
25	10	48.11 ± 1.06	2.19	−0.34
50	10	48.51 ± 0.89	1.84	−1.07
100 ★	10	**48.73 ± 0.72**	1.49	−1.58
200	5	48.79 ± 0.57	1.17	−1.72

300	3	48.57 ± 0.29	0.60	−1.25

Formal convergence on the objective is achieved at N = 300 (CV = 0.60%). The baseline analysis uses N = 100, which does not formally converge on the objective (CV = 1.49%) but remains acceptable for two reasons. First, the portfolio decision, specifically GB capacity, is stable across all sample sizes. Second, the OOS gap at N = 100 (−1.58%) is small in absolute terms and consistent with the negative gaps observed at N = 200 and N = 300, confirming that the in-sample optimism does not grow with the parameter N and is therefore a structural property of the problem rather than a sampling artefact. Results at N = 300 are consistent with the N = 100 solution within 0.4%.

5. Results

We present results in four subsections: (5.1) what deterministic planning gets wrong; (5.2) the cost of tail-risk protection; and (5.3) the value of the stochastic solution and robustness checks.

5.1. What the deterministic model gets wrong

The results for the three models (Risk-neutral (RP), Risk-Averse (RA) and deterministic (EV) are reported in Table 4. The deterministic EV model, optimizing at mean prices (electricity €76.4/MWh, gas €34.7/MWh), selects only a gas boiler (1,499 kW). In contrast, the stochastic model selects a 1,488-kW heat pump paired with a full 1,500 kW gas boiler.

Table. 4. *Investment decisions, annual costs, and carbon performance.*

	RP (β=0)	**RA (β=0.6)**	**EV (det.)**	**EEV**
Installed capacities (kW)				
Heat pump (HP)	1,488	1,453	0	0
Gas boiler (GB)	1,500	1,511	1,499	1,499
Electric boiler (EB)	23.5	270	0	0
Solar thermal (ST)	0	7.0	0	0
TES (kW / kWh)	25 / 101	594 / 2,378	0	0
Annual costs (€/yr)				
Annualized CAPEX	169,053	177,688	16,806	16,940
Fixed O&M	8,989	9,334	4,500	4,501
Expected OPEX	463,078	478,721	—	685,283
Total annual cost (cash)	641,120	665,742	742,087	706,724
Levelized cost and carbon				
LCOH (€/MWh)	48.93	50.80	—	53.93
Carbon (tCO_2/yr)	1,083	1,200	—	2,875
$CVaR_{90}$ of OPEX (€/yr)	—	1,291,301	—	—

The stochastic portfolio replaces the gas-only system with a fuel-flexible hybrid: the HP delivers 78.6% of annual heat demand, the GB supplies 21.1% as backup. This yields LCOH = €48.93/MWh and cuts carbon emissions by 62% (1,083 vs. 2,875 tCO_2/yr). When the deterministic portfolio is evaluated under uncertainty (EEV), annual cost rises to €706,724/yr, a 10.23% penalty driven by crisis exposure that deterministic planning ignores.

5.2. The cost of tail-risk protection

Under risk-averse optimization (β = 0.6, α = 0.90), the portfolio shifts toward larger thermal energy storage (101 → 2,378 kWh) and electric boiler backup (23.5 → 270 kW), at an additional annual cost $24{,}622\ €/yr$ (3.84% above the risk-neutral cash cost). Of this premium, $17{,}195\ €/yr$ falls on normal-year operating cost and only $1{,}674\ €/yr$ on the worst-decile tail, confirming that risk aversion restructures the portfolio rather than reducing peak exposure. Extreme-scenario fuel costs are bounded by crisis-year market prices and cannot be engineered away through investment alone; the larger TES instead allows pre-charging to reduce peak-hour dispatch during crises. The small increase in expected load shedding (5.6 → 197.4 kWh/yr) is a modelling artifact from ramp-rate constraints on the larger TES, representing 0.0015% of annual heat demand with negligible economic impact.

The $CVaR_{90}$ of operating cost is nearly unchanged across risk-aversion levels ($1{,}289{,}627\ €/yr$ at β = 0 vs. $1{,}291{,}301\ €/yr$ at β = 0.6), a finding with direct practical significance: risk aversion changes the portfolio architecture, shifting dispatch optionality into storage, rather than the worst-case outcome itself.

5.3. Value of the stochastic solution and robustness

Table 5 presents the full VSS/EVPI decomposition and scenario method sensitivity. The stochastic solution (RP) costs $641{,}120\ €/yr$ vs. $706{,}724\ €/yr$ for the deterministic portfolio evaluated on the full scenario set (EEV), a VSS of $65{,}604\ €/yr$ (9.28% of EEV, NPV $643{,}987\ €$ over 20 years). The EVPI is $18{,}526\ €/yr$ (2.89% of RP)

VSS is 3.5 times EVPI, indicating that the dominant cost under uncertainty stems from first-stage design decisions, not residual uncertainty. A gas-only portfolio sized at mean prices is over-exposed to joint price crises; the stochastic model corrects this by investing in heat pump capacity upfront. Perfect foresight adds only a further 2.89%.

Robustness to the scenario generation method is confirmed by the Monte Carlo sensitivity: VSS rises to €70,287/yr (11.08% of EEV) under MC sampling, consistent with the LHS result within 2 percentage points, and the HP/GB portfolio is identical across both runs (1,488/1,500 kW vs 1,492/1,500 kW). The LHS method achieves a better dependence fit (Spearman ρ gap 1.9 SEs vs 2.5 SEs at N = 200) and is retained as the baseline method.

Table. 5. *Value of the stochastic solution, EVPI, and scenario method sensitivity.*

Metric	Primary: LHS,	Sensitivity: MC	Interpretation
RP objective (stochastic, β=0)	€641,120/yr	€634,526/yr	Optimal cost under uncertainty
LCOH (RP)	€48.93/MWh	€48.42/MWh	Levelized heat cost
HP / GB capacity	1,488 / 1,500 kW	1,492 / 1,500 kW	Portfolio identical across both runs
EV total cost (deterministic)	€742,087/yr	€732,947/yr	GB-only benchmark
EEV (EV caps on full scenarios)	€706,724/yr	€704,813/yr	Cost of ignoring uncertainty
VSS (= EEV − RP)	**€65,604/yr**	€70,287/yr	Annual benefit of stochastic model
VSS as % of EEV	9.28%	11.08%	Consistent range across both methods
EVPI (= RP − WS)	**€18,526/yr**	—	Value of resolving uncertainty entirely
EVPI as % of RP	2.89%	—	VSS is 3.5× EVPI — most value from modelling, not resolving
NPV of VSS (20 yr, 8%)	€643,987	€690,060	Present-value planning savings
Spearman ρ gap (target +0.964)	1.9 SEs (N=100)	2.5 SEs (N=200)	LHS achieves better dependence fit
Out-of-sample obj. (N=500)	€637,120/yr	—	Gap vs. in-sample: −0.62%

Despite the attenuation of dependence (ρ = 0.773 vs. 0.964), the optimal portfolio remains unchanged, indicating that the investment result is robust to moderate variations in cross-commodity correlation

6. Conclusion

The choice of independent commodity price dynamics is not neutral: it systematically underestimates risk and changes the technology portfolio an industrial heat operator should invest in. By combining jump-driven price processes, a 4-dimensional *t*-copula, and a mean–CVaR two-stage stochastic LP, the proposed framework selects a qualitatively different relative to deterministic planning.

The VSS of $65{,}604\ €/yr$ (9.28% of EEV, NPV $643{,}987\ €$ over 20 years) is 3.5 times the EVPI, indicating that the dominant cost under uncertainty stems from first-stage design decisions rather than residual uncertainty. A gas-only portfolio sized at mean prices cannot be recovered from once a joint price crisis materializes; the stochastic model captures this irreversibility by committing to heat pump capacity upfront.

Risk-averse optimization (β = 0.6) adds thermal energy storage (2,378 kWh) and electric boiler backup (270 kW) for a 3.84% cost premium, yet $CVaR_{90}$ changes by only €1,674/yr. Extreme-scenario fuel costs are bounded by crisis-year market prices and cannot be engineered away; risk aversion shifts portfolio architecture rather than the worst-case outcome itself.

Three simplifications bias results conservatively: fixed thermal demand, a static carbon price at $€65/tCO_2$, and a discrete year pool capping Spearman correlation at 0.773 (target 0.964). Immediate extensions include stochastic carbon pricing, demand-side flexibility, and scaling to multi-facility district heat networks.

Acknowledgements

This work is supported by the French Ministry of Higher Education and Research through a doctoral contract at INSA Lyon (CETHIL, UMR CNRS 5008).

Nomenclature

β	CVaR risk-aversion weight (–)	C_{HP}	Heat pump capacity (kW)
α	CVaR confidence level (–)	C_{GB}	Gas boiler capacity (kW)
η_2	Second-law HP efficiency (–)	C_{ST}	Solar thermal capacity (kW)
η_{opt}	Solar collector optical eff. (–)	C_{EB}	Electric boiler capacity (kW)
U_L	Collector heat loss coeff. (W/m²K)	C_{TES}	TES capacity (kW / kWh)
COP	Coefficient of performance (–)	LCOH	Levelized cost of heat (€/MWh)
CRF	Capital recovery factor (–)	VSS	Value of stochastic solution (€)

References

[1] M. Rehfeldt, C. Rohde, T. Fleiter, F. Toro, and F. Reitze, “A bottom-up estimation of heating and cooling demand in the European industry,” *Eceee Ind. Effic. 2016 Proc.*, vol. 11, Jun. 2018, doi: 10.1007/s12053-017-9571-y.

[2] “REPowerEU.” Accessed: Mar. 17, 2026. [Online]. Available: https://commission.europa.eu/topics/energy/repowereu_en

[3] A. K. Dixit and R. S. Pindyck, *Investment under Uncertainty*. Princeton University Press, 1994. doi: 10.2307/j.ctt7sncv.

[4] J. van Ouwerkerk *et al.*, “Quantifying benefits of renewable investments for German residential Prosumers in times of volatile energy markets,” *Nat. Commun.*, vol. 15, no. 1, p. 8206, Sep. 2024, doi: 10.1038/s41467-024-51967-6.

[5] “Industrial heat electrification: New business models emerge | McKinsey & Company.” Accessed: Mar. 17, 2026. [Online]. Available: https://www.mckinsey.com/industries/energy-and-materials/our-insights/blog/industrial-heat-electrification-in-europe-new-business-models-emerge?utm_medium=DSMN8&utm_source=LinkedIn&utm_user=14419233756665344

[6] H. Lund, N. Duic, P. A. Østergaard, and B. V. Mathiesen, “Future district heating systems and technologies: On the role of smart energy systems and 4th generation district heating,” *Energy*, vol. 165, pp. 614–619, Dec. 2018, doi: 10.1016/j.energy.2018.09.115.

[7] G. Mavromatidis, K. Orehounig, and J. Carmeliet, “Design of distributed energy systems under uncertainty: A two-stage stochastic programming approach,” *Appl. Energy*, vol. 222, pp. 932–950, Jul. 2018, doi: 10.1016/j.apenergy.2018.04.019.

[8] M. A. Lozano, J. C. Ramos, and L. M. Serra, “Cost optimization of the design of CHCP (combined heat, cooling and power) systems under legal constraints,” *Energy*, vol. 35, no. 2, pp. 794–805, Feb. 2010, doi: 10.1016/j.energy.2009.08.022.

[9] R. M. Lima, A. J. Conejo, L. Giraldi, O. Le Maître, I. Hoteit, and O. M. Knio, “Risk-Averse Stochastic Programming vs. Adaptive Robust Optimization: A Virtual Power Plant Application,” *Inf. J. Comput.*, vol. 34, no. 3, pp. 1795–1818, May 2022, doi: 10.1287/ijoc.2022.1157.

[10] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. in Springer Series in Operations Research and Financial Engineering. New York, NY: Springer, 2011. doi: 10.1007/978-1-4614-0237-4.

[11] M. C. Fernandes, J. C. Dias, and J. P. V. Nunes, “Modeling energy prices under energy transition: A novel stochastic-copula approach,” *Econ. Model.*, vol. 105, p. 105671, Dec. 2021, doi: 10.1016/j.econmod.2021.105671.

[12] F. Durante, A. Gianfreda, F. Ravazzolo, and L. Rossini, “A Multivariate Dependence Analysis for Electricity Prices, Demand and Renewable Energy Sources,” Jan. 04, 2022, *arXiv*: arXiv:2201.01132. doi: 10.48550/arXiv.2201.01132.

[13] “CEPA study on the Value of Lost Load in the electricity supply.pdf.” Accessed: Mar. 10, 2026. [Online]. Available: https://www.acer.europa.eu/sites/default/files/documents/en/Electricity/Infrastructure_and_network%20

development/Infrastructure/Documents/CEPA%20study%20on%20the%20Value%20of%20Lost%20Load%20in%20the%20electricity%20supply.pdf

[14] ADEME, “Base carbone®.” Accessed: Mar. 10, 2026. [Online]. Available: https://data.ademe.fr/datasets/undefined

[15] “Bilan électrique 2023 - Emissions | RTE.” Accessed: Mar. 17, 2026. [Online]. Available: https://analysesetdonnees.rte-france.com/bilan-electrique-2023/emissions?utm_source=chatgpt.com#Introduction

[16] “EIB Project Carbon Footprint Methodologies,” 2023.

[17] “2024 Carbon Market Report: a stable and well-functioning market, driving emissions from power and industry installations to a historic reduction of 16.5% - Climate Action.” Accessed: Mar. 17, 2026. [Online]. Available: https://climate.ec.europa.eu/news-other-reads/news/2024-carbon-market-report-stable-and-well-functioning-market-driving-emissions-power-and-industry-2024-11-19_en

[18] I. Staffell, D. Brett, N. Brandon, and A. Hawkes, “A review of domestic heat pumps,” *Energy Environ. Sci.*, vol. 5, no. 11, pp. 9291–9306, Oct. 2012, doi: 10.1039/C2EE22653G.

[19] I. Tsiropoulos, D. Tarvydas, and A. Zucker, “Cost development of low carbon energy technologies: Scenario-based cost trajectories to 2050, 2017 edition,” JRC Publications Repository. Accessed: Mar. 10, 2026. [Online]. Available: https://publications.jrc.ec.europa.eu/repository/handle/JRC109894

[20] J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*, 1st ed. Wiley, 2013. doi: 10.1002/9781118671603.

[21] P. Artzner, F. Delbaen, J. Eber, and D. Heath, “Coherent Measures of Risk,” *Math. Finance*, vol. 9, no. 3, pp. 203–228, Jul. 1999, doi: 10.1111/1467-9965.00068.

[22] R. T. Rockafellar and S. Uryasev, “Optimization of conditional value-at-risk,” *J. Risk*, vol. 2, no. 3, pp. 21–41, 2000, doi: 10.21314/JOR.2000.038.

[23] T. Meyer-Brandis and P. Tankov, “MULTI-FACTOR JUMP-DIFFUSION MODELS OF ELECTRICITY PRICES,” *Int. J. Theor. Appl. Finance*, vol. 11, no. 05, pp. 503–528, Aug. 2008, doi: 10.1142/S0219024908004907.

[24] H. Geman and A. Roncoroni, “Understanding the Fine Structure of Electricity Prices,” *J. Bus.*, vol. 79, no. 3, pp. 1225–1261, 2006, doi: 10.1086/500675.

[25] R. C. Merton, “Option pricing when underlying stock returns are discontinuous,” *J. Financ. Econ.*, vol. 3, no. 1, pp. 125–144, Jan. 1976, doi: 10.1016/0304-405X(76)90022-2.

[26] S. G. Kou, “A Jump-Diffusion Model for Option Pricing,” *Manag. Sci.*, vol. 48, no. 8, pp. 1086–1101, 2002.

[27] “Sklar, A. (1959) Fonctions de Répartition à n Dimensions et Leurs Marges. Publications de l’Institut Statistique de l’Université de Paris, 8, 229-231. - References - Scientific Research Publishing.” Accessed: Mar. 30, 2026. [Online]. Available: https://www.scirp.org/reference/ReferencesPapers?ReferenceID=%201834923

[28] S. Demarta and A. J. McNeil, “The t Copula and Related Copulas,” *Int. Stat. Rev.*, vol. 73, no. 1, pp. 111–129, Jan. 2007, doi: 10.1111/j.1751-5823.2005.tb00254.x.

[29] “Key developments in European electricity and gas markets - 2025 ACER Monitoring Report.” Accessed: Mar. 27, 2026. [Online]. Available: https://www.acer.europa.eu/sites/default/files/documents/Publications/2025_ACER_Gas_Electricity_Key_Developments.pdf

[30] P. Kotek, A. Selei, B. Takácsné Tóth, and B. Felsmann, “What can the EU do to address the high natural gas prices?,” *Energy Policy*, vol. 173, no. C, 2023, Accessed: Mar. 27, 2026. [Online]. Available: https://ideas.repec.org//a/eee/enepol/v173y2023ics0301421522005316.html

[31] M. Mckay, R. Beckman, and W. Conover, “A Comparison of Three Methods for Selecting Vales of Input Variables in the Analysis of Output From a Computer Code,” *Technometrics*, vol. 21, pp. 239–245, May 1979, doi: 10.1080/00401706.1979.10489755.

[32] M. Kaut, H. Vladimirou, S. W. Wallace, and S. Zenios, “Stability analysis of portfolio management with conditional value-at-risk,” *Quant. Finance*, vol. 7, no. 4, pp. 397–409, 2007, doi: 10.1080/14697680701483222.