An Affine Arithmetic based Methodology for Energy Hub Operation-Scheduling in the presence of Data Uncertainty

A. Vaccaro(1), C. Pisani(1), A.F. Zobaa(2)

(1)University of Sannio, Department of Engineering, Piazza Roma 21, 82100 Benevento, Italy, vaccaro@unisannio.it
(1) Brunel University, Department of Electronic and Computer Engineering, Howell Building, Brunel University London, Uxbridge, UB8 3PH, United Kingdom

Abstract
In this paper the role of self-validated computing for solving the energy hub-scheduling problem in the presence of multiple and heterogeneous sources of data uncertainties is explored and a new solution paradigm based on Affine Arithmetic (AA) is conceptualized. The benefits deriving from the application of this methodology are analysed in details, and several numerical results are presented and discussed.

1.0 Introduction

Traditional energy networks are today subject to many demanding challenges such as aging infrastructures, need for new generation facilities, network expansion to meet growing energy demand, distributed energy resources and reliability coordination.

In this complex scenario, the large-scale deployment of the Energy Hub paradigm represents one of the most promising enabling technology aimed at supporting the evolution of traditional energy networks according to holistic, proactive, reconfigurable and self-healing web-based architectures based on distributed, self-organizing and cooperative energy resources.

From a conceptual point of view, the energy hub can be considered as a set of interconnected converter and storage systems aimed at processing multiple energy carriers and providing high-value energy services at its load ports (e.g. electricity, heating) [1-3]. A wide and heterogeneous spectrum of technologies can be adopted to implement an energy hub, including combined heat and power technology, tri-generation systems, power-electronic devices, and heat exchangers. These enabling technologies allow the energy hub to supply its loads by means of multiple and redundant paths, characterized by different combinations of energy carriers, which introduce many degree of freedoms in energy hub supply compared to the traditional, decoupled energy systems [4,5].

This greater supply flexibility, if properly managed, could increase the overall efficiency of energy networks by: (i) supporting massive pervasion of distributed generation and storage energy systems; (ii) facilitating the integration of renewable energy sources; (iii) reducing system losses and greenhouse gas emissions; (iv) increasing the reliability of the energy supply to the customers [6,7]. Consequently, a significant growth in the number of geographically dispersed energy hubs connected to energy networks is expected in the near future.

From this perspective, a crucial issue is how to attain a reliable and cost effective operation of the interconnected energy hubs by properly dispatching their input energy carriers, which could be characterized by different features such as cost, availability, reliability and environmental impacts.

This operation-scheduling problem could be formalized by a non-linear constrained optimization problem, whose objective function measures the energy scheduling effectiveness and the problem constraints include both equality (e.g. the energy balance in the energy hub) and inequality (e.g. energy converter ratings) functions.

Solution methods traditionally adopted to solve this problem generally assume that both the input data and the hub parameters are described by crisp values, which should be fixed by the analyst based on preliminary studies and simplified hypothesis about the energy hub under study. These approaches, here referred as deterministic dispatch algorithms, identify operation-scheduling strategies, which are deemed effective for a limited set of energy hub operation states. Thus, in the presence of data uncertainties, comprehensive scenario analysis, aimed at identifying the correct solution domain, should be implemented.

This is a major issue in real world applications, where multiple and heterogeneous source of uncertainties could sensibly affect the reliable and cost effective energy hub operation, such as the complex dynamics of the energy costs, the random energy demand fluctuations, the uncertain power injections from renewable power generators and the non-idealities characterizing the operation of the energy hub elements. In this context, the research for reliable solution methodologies aimed at solving the energy hub operation-scheduling problem in the presence of uncertain data represents one of the most relevant issues to address [8,9]. The adoption of these methodologies allows the analyst to compute both the data tolerance, i.e. uncertainties characterization, and the solution tolerance, i.e. uncertainty propagation, providing insight into the confidence level of the operation-scheduling strategy. Moreover, it can effectively support the analyst in computing comprehensive sensitivity analysis aimed at estimating the rate of change in the problem solution with respect to large changes of the input data.

Conventional methodologies available in the literature address this problem by means of sampling and analytical techniques [8,9], which aim at modelling the variability and stochastic nature of the uncertain data. In particular, uncertainty analysis based on sampling based methods require several model runs aimed at sampling various combinations of the uncertain data, and, since the number of simulations may be rather large, especially in the presence of large and uncorrelated uncertain parameters [20,21], the required computational burden could be prohibitively expensive [10].

Analytical techniques are computationally more effective, but they require some mathematical assumptions in order to simplify the problem and obtaining an effective characterization of the output random variables [18]. These assumptions are typically based on convolution techniques and fast Fourier transform. However, as discussed in [11,12], analytical techniques present various shortcomings, such as the statistical dependence of the input data, and the problems associated with accurately identifying probability distributions for some input data [22,23]. These issues are not infrequent in energy hub analysis, since the analyst is not always confident in translating its imprecise knowledge in terms of probability distributions for some input variables, such as the power generated by renewable power generators, due to his/her qualitative knowledge and the lack of sufficient data. To face this issue, the analyst makes often the assumption of normality and statistical independence of the input data, but experimental results show that these assumptions are often not supported by empirical evidence. These drawbacks may limit the use of analytical methods in practical applications, especially for the study of complex energy hubs [24].

In order to overcome some of the aforementioned limitations of sampling and analytical methods, more sophisticated techniques for uncertainty analysis of complex systems, based on the self-validated computing theory, have been recently proposed in the literature [25,26]. The main advantage of this paradigm is that it keeps track of the accuracy of the computed quantities, as part of the process of computing them, without requiring information about the type of uncertainty in the parameters [25]. The simplest and most popular of these models is Interval Mathematics (IM), which allows for numerical computation where each quantity is represented by an interval of floating point numbers without a probability structure [26]. However, the adoption of this solution technique present many drawbacks derived mainly by the so called "dependency problem" and "wrapping effect" [25, 27], which make the solution provided by an IM method not always as informative as expected.

To overcome these limitations, in this paper a more effective self-validated paradigm based on Affine Arithmetic (AA) is proposed to solve the energy hub operation-scheduling problem [13,14]. AA is an instance of self-validated computing in which all the problem variables are represented by means of affine combinations of certain primitive variables [15]. The main benefits of AA is that, unlike standard IM, it propagates the uncertainties by keeping track of their correlations, and, consequently, it is less affected by the loss of precision often observed in long interval computations [16].

The application of AA for uncertainty representation in energy hub operation scheduling allows the analyst to express each decision variable by a central value and a set of partial deviations. These deviations are associated with as many noise variables as those describing the effect of the various uncertainty sources affecting the energy hub operation.

The parameters of these affine forms can be computed by approximating the objective function and the equality and inequality constraints describing the operation-scheduling problem by a linear relaxation based on AA. Starting from this set of equations, the energy hub operation-scheduling problem in the presence of data uncertainty can be formalised by a linear multiobjective programming problem. The solution of this problem allows the analyst to assess an enclosure of the objective function range, which is guaranteed to contain its actual value. Moreover, the "nominal" energy hub dispatching strategy can be defined according to the central value of the affine forms of the decision variables, while the corresponding partial deviations can be used to correct the dispatching strategy depending by the actual energy hub operation point. These important features allows the analyst to solve simultaneously both the short/medium (i.e. one day ahead) and the real time operation-scheduling problem.

2.0 Problem formulation

2.1 Mathematical background

In energy hub dispatch analysis, the energy hub is typically modelled by a multi-inputs/multi-outputs system aimed at converting the input energy flows $\mathbf{E}=\left[E_1,..,E_N\right]$, into the output energy flows $\mathbf{L}=\left[L_1,..,L_M\right]$ [2,8]. Depending by the particular architecture deployed, each input energy flow E_i can feed n_i internal converters according to certain dispatching strategies, which are described by the vector $\left(E_i^1,..,E_i^{n_i}\right)$ with $E_i=\sum_{j=1}^{n_i} E_i^j$. The internal converters are assumed to operate in steady state, and their conversion efficiencies are modelled by static algebraic equations that, for the j^{th} converter, assume the following structure:

$$E_k^{ij}=\eta_{i,k}^j E_i^j \quad (1)$$

where $\eta_{i,k}^j$ is the nameplate efficiency in converting the input energy carrier i to the output energy carrier k, while E_i^j and E_k^{ij} are the input and output energy flow, respectively. These energy flows can be combined, stored and/or processed by other energy converters in order to supply the energy hub loads $\mathbf{L}$. Consequently, the resulting energy hub model can be described by the following equations:

$$\begin{bmatrix} L_1 & \dots & L_M \end{bmatrix}^T = \mathbf{A} \begin{bmatrix} E_1^1 & \dots & E_1^{n_1} & E_2^1 & \dots & E_2^{n_2} & \dots & E_N^1 & \dots & E_N^{n_N} \end{bmatrix}^T \quad (2)$$

where the matrix $\mathbf{A}$ is called the converter coupling matrix, whose elements could be zeros, converter efficiencies or product of converter efficiencies. This mathematical formalization can be easily extended to model the energy storage systems, as detailed described in [2,8,9].

Analyzing the model formalized in (2), it is worth observing as the energy hub load requirements can be satisfied by means of multiple energy carriers and different combinations of them. Consequently, a proper strategy aimed at identifying the most effective energy dispatching strategy is required for a reliable and cost effective energy hub operation. In particular, if the energy hub does not integrate energy storage systems, the overall problem can be formalized by the following constrained nonlinear programming problem:

$$\begin{cases} \min\limits_{\mathbf{x}} \quad f(\mathbf{x}) \\ s.t. \quad g_i(\mathbf{x}) = 0 \quad i = 1,..,n_g \\ \qquad h_j(\mathbf{x}) < 0 \quad j = 1,..,n_h \end{cases} \quad (3)$$

Where, n_g is the number of equality constraints, n_h is the number of inequality constraints and $\mathbf{x}$ is the vector of the decision variables defining the energy hub dispatching strategy, namely:

$$\mathbf{x} = \begin{bmatrix} E_1^1 & \dots & E_1^{n_1} & E_2^1 & \dots & E_2^{n_2} & \dots & E_N^1 & \dots & E_N^{n_N} \end{bmatrix}^T \quad (4)$$

The objective function $f(\mathbf{x})$ could integrate both technical and economic criteria including the minimization of the hourly energy cost, the minimization of the loss of load probability, the maximization of the renewable power exploitation etc.

The inequality constraints include the minimum and maximum allowable limits for each energy converter, the maximum allowable limit for each input energy carrier (i.e., $E_i \le E_{\max,i} \quad i = 1,..,N$) and the minimum and maximum allowable limits for each decision variable (namely $x_{\min,i} \le x_i \le x_{\max,i} \quad i = 1,..,\sum_{j=1}^{N} n_j$).

In addition, the decision variables should satisfy the energy hub equations (2), which represent the equality constraints for problem (3).

If the energy hub integrates storage systems, the decision variables should include the quantity of stored energy for each time period. In this case, the problem (3) should be solved over multiple time periods in order to provide decisions on energy flows to be purchased and stored at each point in time.

2.2 Sources of Uncertainty in Energy Hub Operation Scheduling

The mathematical formulation described in the previous section has all input data specified from the snapshot corresponding to a point in time or from a proper set of "crisp" values, e.g. the expected generation/load profiles, which should be fixed for the energy hub under study. Consequently, the corresponding solution of the operation scheduling problem is deemed representative of a limited set of energy hub operation states and, if the input conditions are uncertain, numerous scenarios need to be analysed in order to cover the entire uncertainty range.

Numerous sources, both internal and external to the system, generate uncertainties in energy hub operation. Many uncertainties derive from the unpredictable dynamics of the energy prices in both the spot markets, which depend on the specific market structure and are typically characterised by high volatility, and the bilateral markets, where the parties involved fix the prices spread depending on the corresponding spot prices. In both cases, the effects of imbalance penalties, deriving by the deviations from the load/generation schedules, should be accurately considered, since it could compromise the cost effectiveness of the operation scheduling strategy.

Other relevant uncertainties derive by the random dynamics of the energy hub loads, which are influenced by several external factors such as economic, season and weather effects. Forecasting these complex dynamics involves large uncertainty, especially on medium and long-term time scenario.

Renewable power generators (e.g. wind, solar) induce further uncertainties, since the corresponding generated power profiles varies over time following the natural fluctuations of their energy sources.

In particular, experimental studies have shown that the power generated by solar systems is highly influenced by the random clouds coverage, which makes short-term solar energy forecasting a very difficult task [28]. In addition, short-term and hourly fluctuations of wind energy systems are hard to predict, although their generation profiles may follow a generally well-known daily or seasonal pattern [29].

Finally, the approximations errors due to the application of simplified algebraic equations in modelling the internal energy converters, cause further uncertainties and complex correlations in uncertainty propagation.

All these uncertainties complicate the mathematical formulation of the energy hub dispatch problem by introducing a lack of determinism in the problem solution. In this condition, the overall problem can be formalized by the following constrained uncertain optimization problem:

$$\begin{cases} \min\limits_{\tilde{\mathbf{x}}} \quad f(\tilde{\mathbf{x}},\tilde{\mathbf{p}}) \\ s.t. \quad g_i(\tilde{\mathbf{x}},\tilde{\mathbf{p}}) = 0 \quad i = 1,..,n_g \\ \qquad h_j(\tilde{\mathbf{x}},\tilde{\mathbf{p}}) < 0 \quad j = 1,..,n_h \end{cases} \quad (5)$$

where $\tilde{\mathbf{p}}$ is the vector of the uncertain parameters. To solve this uncertain optimization problem, it is necessary to design effective computational paradigms for:

1. representing the vector of the uncertainties parameters (i.e. intervals, fuzzy numbers),
2. processing the data uncertainties by defining proper mathematical operators (i.e. aimed at computing $\tilde{y} = f(\tilde{\mathbf{x}},\tilde{\mathbf{p}})$),
3. checking the consistency of equality and inequality constraints by defining proper relational operators between uncertain variables (i.e. $\tilde{y}_1 \le \tilde{y}_2$, $\tilde{y}_1 = \tilde{y}_2$),
4. solving linear and non-linear system of equations in the presence of data uncertainties (i.e. find $\tilde{\mathbf{x}}\ ''\ f(\tilde{\mathbf{x}},\tilde{\mathbf{p}}) = 0$, where $\tilde{\mathbf{p}}$ is a fixed uncertain vector).

To address these issues, in this paper an Affine Arithmetic based framework is conceptualized.

3.0 Energy Hub Operation Scheduling by Affine Arithmetic

3.1 Mathematical Preliminaries

Affine Arithmetic (AA) is a range-based formalism for numerical computation, which allows to represent heterogeneous uncertainty sources both external and internal to the system under study [14,15]. In AA, each variable x is represented by an affine form $\hat{x}$, which is a first-degree polynomial of the form [16,17]:

$$\hat{x} = x_0 + x_1\varepsilon_1 + x_2\varepsilon_2 + .. + x_n\varepsilon_n \quad (6)$$

Where, x_0 is the central value of the affine form $\hat{x}$ and $x_i \in \Re$ are the corresponding partial deviations. The variables ε_i, which are assumed to be “unknown but bounded” in the interval [-1,1], are the “noise symbols” of the affine form $\hat{x}$ and may be shared between all the affine forms involved in the computing process.

Affine forms may be processed by means of proper AA based mathematical operators, obtained by replacing the elementary real-number operators with the corresponding affine versions. This process is straightforward for linear mappings, since the corresponding affine extension can be obtained by expanding and rearranging only the noise symbols characterizing the affine forms $\hat{x}$ and $\hat{y}$:

$$\begin{aligned}\hat{x} \pm \hat{y} = (x_0 \pm y_0) + (x_1 \pm y_1)\varepsilon_1 + (x_2 \pm y_2)\varepsilon_2 + \\ .. + (x_n \pm y_n)\varepsilon_n\end{aligned} \quad (7)$$

$$\alpha\hat{x} = (\alpha x_0) + (\alpha x_1)\varepsilon_1 + (\alpha x_2)\varepsilon_2 + .. + (\alpha x_n)\varepsilon_n \quad \forall \alpha \in \mathrm{R} \ (8)$$

$$\hat{x} \pm \lambda = (x_0 \pm \lambda) + x_1\varepsilon_1 + x_2\varepsilon_2 + .. + x_n\varepsilon_n \qquad \forall \lambda \in \mathrm{R} \ (9)$$

On the other hand, if f is a nonlinear function, the corresponding affine extension cannot be described by an affine combination of the noise symbols ε_i:

$$\begin{aligned}\hat{z} = f\left(\hat{x}, \hat{y}\right) = f(x_0 + x_1\varepsilon_1 + x_2\varepsilon_2 + .. + x_n\varepsilon_n, y_0 + \\ + y_1\varepsilon_1 + y_2\varepsilon_2 + .. + y_n\varepsilon_n) = f^*(\varepsilon_1, \varepsilon_2, .., \varepsilon_n)\end{aligned} \quad (10)$$

Therefore, in this case it is necessary to identify an affine function, which approximates the function $f^*(\varepsilon_1, \varepsilon_2, .., \varepsilon_n)$ over its domain:

$$f^a(\varepsilon_1, \varepsilon_2, .., \varepsilon_n) = z_0 + z_1\varepsilon_1 + ... + z_n\varepsilon_n \ (11)$$

and bound the corresponding approximation error:

$$\hat{z} = f^a(\varepsilon_1, \varepsilon_2, .., \varepsilon_n) + z_{n+1}\varepsilon_{n+1} = \\ = z_0 + z_1\varepsilon_1 + ... + z_n\varepsilon_n + z_{n+1}\varepsilon_{n+1} \quad (12)$$

where the term $z_{n+1}\varepsilon_{n+1}$ represents the approximation error defined as:

$$e^*(\varepsilon_1, \varepsilon_2, .., \varepsilon_n) = f^*(\varepsilon_1, \varepsilon_2, .., \varepsilon_n) - f^a(\varepsilon_1, \varepsilon_2, .., \varepsilon_n) \quad (13)$$

The new noise ε_{n+1} introduced in (12) is distinct from all other ε_i, and the corresponding partial deviation $z_{n+1} \in \Re^+$ is an upper bound of the absolute magnitude of e^*.

The approximation function f^a could be defined as follows:

$$f^a(\varepsilon_1, \varepsilon_2, .., \varepsilon_n) = \alpha\hat{x} + \beta\hat{y} + \varsigma \quad (14)$$

where the unknown coefficients α, β and ζ can be identified according to the Chebyshev's approximation theorem for univariate functions [15].

The described computing paradigm can be generalized in order to define affine form approximations of any N -dimensional real value function $f(x_1, .., x_N)$:

$$\hat{f} = f_0 + \sum_{i=1}^{n} f_i\varepsilon_i + f_{n+1}\varepsilon_{n+1} \quad (15)$$

where the affine form $\hat{f}$ is by definition an inclusion function:

$$\forall(x_1, .., x_N) \in X^N, f(x_1, .., x_N) \in f_0 + \sum_{i=1}^{n} f_i\varepsilon_i + f_{n+1}\varepsilon_{n+1} \quad (16)$$

This important result leads to the following propositions [18]:

$$\text{If } \forall(x_1, .., x_N) \in X^N, f(x_1, .., x_N) \leq 0 \text{ then } \forall(y_1, .., y_n) \in [-1,1]^n, \sum_{i=1}^{n} f_i y_i \leq f_{n+1} - f_0 \quad (17)$$

$$\text{If } \forall(x_1, .., x_N) \in X^N, f(x_1, .., x_N) = 0 \text{ then } \forall(y_1, .., y_n) \in [-1,1]^n, \begin{cases} \sum_{i=1}^{n} f_i y_i \leq f_{n+1} - f_0 \\ -\sum_{i=1}^{n} f_i y_i \leq f_{n+1} + f_0 \end{cases} \quad (18)$$

3.2 The AA based Solution Paradigm

AA can be effectively adopted for uncertainty representation in energy hub operation scheduling. According to this paradigm, each decision variable is expressed by a central value and a set of partial deviations. These deviations are associated with as many noise variables as those which describe the effect of the various phenomena affecting the energy hub operation. Without loss of generality the uncertainties considered here are those associated with the hub loads, the energy cost prices and the electrical energy produced by renewable generators. Therefore, the input variables can be expressed by the following affine forms:

$$\begin{cases} \hat{L}_j = L_{j,0} + L_{j,1}\varepsilon_{L_j} \quad j = 1,..,M \\ c\hat{E}_j = cE_{j,0} + cE_{j,1}\varepsilon_{cE_j} \quad j = 1,..,N \\ \hat{E}r_j = Er_{j,0} + Er_{j,1}\varepsilon_{Er_j} \quad j = 1,..,N_g \end{cases} \quad (19)$$

Where:

- $L_{j,0}$ and $L_{j,1}$ are the central value and the partial deviation of the $j-th$ hub load respectively;
- $cE_{j,0}$ and $cE_{j,1}$ are the central value and the partial deviation of the $j-th$ energy carrier cost respectively;
- $Er_{j,0}$ and $Er_{j,1}$ are the central value and the partial deviation of the electrical energy produced by the $j-th$ renewable generator respectively;
- ε_{L_j} is the noise symbol representing the uncertainty affecting the energy demanded by the $j-th$ hub load;
- ε_{cE_j} is the noise symbol representing the uncertainty affecting the $j-th$ energy carrier cost;
- ε_{Er_j} is the noise symbol representing the uncertainty affecting the electrical energy produced by the $j-th$ renewable power generator.

Consequently, the affine forms representing the decision variables are:

$$\hat{x}_i = x_0^i + \sum_{j=1}^{M} x_{L_j}^i \varepsilon_{L_j} + \sum_{j=1}^{N} x_{cE_j}^i \varepsilon_{cE_j} + \sum_{j=1}^{N_g} x_{Er_j}^i \varepsilon_{Er_j} \quad i = 1,..,\sum_{j=1}^{N} n_j \quad (20)$$

where:

- x_0^i is the central value of the i-th decision variable;
- $x_{L_j}^i$ is the partial deviation of the i-th decision variable due to the j-th hub load;
- $x_{cE_j}^i$ is the partial deviation of the i-th decision variable due to the j-th energy carrier cost ;
- $x_{Er_j}^i$ is the partial deviation of the i-th decision variable due to the electrical energy produced by the j-th renewable power generator;

The parameters of these affine forms can be computed by approximating the mathematical programming problem (3) by a linear relaxation. The solution of this linear programming problem yields bounds or a certificate of unfeasibility of the original hub operation scheduling problem. One of the originality of this approach lies in how the linear relaxation is made. In particular we propose a linear approximation of each equation of the problem (3) based on AA.

In particular, according to the AA computing paradigm, each inequality constraint is relaxed by one linear equation and each equality constraint by two linear equations:

$$\begin{gathered}\hat{h}_k = h_0^k + \sum_{j=1}^{M} h_{L_j}^k \varepsilon_{L_j} + \sum_{j=1}^{N} h_{cE_j}^k \varepsilon_{cE_j} + \sum_{j=1}^{N_g} h_{Er_j}^k \varepsilon_{Er_j} + h_{n_x+1}^k \varepsilon_{n_x+1} \le 0 \Rightarrow \\ \sum_{j=1}^{M} h_{L_j}^k \varepsilon_{L_j} + \sum_{j=1}^{N} h_{cE_j}^k \varepsilon_{cE_j} + \sum_{j=1}^{N_g} h_{Er_j}^k \varepsilon_{Er_j} \le h_{n_x+1}^k - h_0^k \qquad k = 1,..,n_h\end{gathered} \quad (21)$$

$$\begin{gathered}\hat{g}_i = g_0^i + \sum_{j=1}^{M} g_{L_j}^i \varepsilon_{L_j} + \sum_{j=1}^{N} g_{cE_j}^i \varepsilon_{cE_j} + \sum_{j=1}^{N_g} g_{Er_j}^i \varepsilon_{Er_j} + g_{n_x+1}^i \varepsilon_{n_x+1} = 0 \Rightarrow \\ \begin{cases} \sum_{j=1}^{M} g_{L_j}^i \varepsilon_{L_j} + \sum_{j=1}^{N} g_{cE_j}^i \varepsilon_{cE_j} + \sum_{j=1}^{N_g} g_{Er_j}^i \varepsilon_{Er_j} \le g_{n_x+1}^i - g_0^i \\ -\sum_{j=1}^{M} g_{L_j}^i \varepsilon_{L_j} + \sum_{j=1}^{N} g_{cE_j}^i \varepsilon_{cE_j} + \sum_{j=1}^{N_g} g_{Er_j}^i \varepsilon_{Er_j} \le g_{n_x+1}^i + g_0^i \end{cases} \qquad i = 1,..,n_g\end{gathered} \quad (22)$$

Where $n_x = M + N + N_g$ is the number of noise symbols describing the energy hub uncertainty while $\hat{h}_k$ and $\hat{g}_i$ are the affine approximation of $h_k(\mathbf{x})$ and $g_i(\mathbf{x})$ respectively. Moreover, it is important to note the presence of the new noise symbol ε_{n_x+1}, which represents the uncertainty due to the affine approximation errors.

Following the same approach we can compute the affine approximation of the objective function $f(\mathbf{x})$:

$$\hat{f} = f_0 + \sum_{j=1}^{M} f_{L_j}\varepsilon_{L_j} + \sum_{j=1}^{N} f_{cE_j}\varepsilon_{cE_j} + \sum_{j=1}^{N_g} f_{Er_j}\varepsilon_{Er_j} + f_{n_x+1}\varepsilon_{n_x+1} \quad (23)$$

Staring form this affine function it is possible to define several criteria aimed at describing the risk attitude of the analyst. For example a worst case solution can be identified by minimizing the upper bound of $\hat{f}$, a reliable solution can be identified by minimizing the radius of $\hat{f}$, an "average" solution can be identified by minimizing the central value of $\hat{f}$. In our own opinion a suitable trade off between the "reliable" and the "average" criteria represents the most attractive solution.

Following this direction, the energy hub operation-scheduling problem in the presence of data uncertainty can be formalized according to the following linear multiobjective programming problem:

$$\min_{\mathbf{y}} \left(f_0, rad(\hat{f}) \right)$$

$$\begin{cases} \sum_{j=1}^{M} g^i_{L_j}\varepsilon_{L_j} + \sum_{j=1}^{N} g^i_{cE_j}\varepsilon_{cE_j} + \sum_{j=1}^{N_g} g^i_{Er_j}\varepsilon_{Er_j} \le g^i_{n_x+1} - g^i_0 \\ -\sum_{j=1}^{M} g^i_{L_j}\varepsilon_{L_j} + \sum_{j=1}^{N} g^i_{cE_j}\varepsilon_{cE_j} + \sum_{j=1}^{N_g} g^i_{Er_j}\varepsilon_{Er_j} \le g^i_{n_x+1} + g^i_0 \end{cases} \quad i = 1,..,n_g \quad (24)$$

$$\sum_{j=1}^{M} h^k_{L_j}\varepsilon_{L_j} + \sum_{j=1}^{N} h^k_{cE_j}\varepsilon_{cE_j} + \sum_{j=1}^{N_g} h^k_{Er_j}\varepsilon_{Er_j} \le h^k_{n_x+1} - h^k_0 \qquad k = 1,..,n_h$$

Where $\mathbf{y} = \left[\mathbf{y}_1^T,..,\mathbf{y}_{n_Y}^T \right]$ is the decision variables vector defined as follows:

$$\mathbf{y}_i = \left[x^i_0, x^i_{L_1},..x^i_{L_M}, x^i_{cE_1},..,x^i_{cE_N}, x^i_{Er_1},..,x^i_{Er_{Ng}} \right] \quad i = 1,..,n_Y = \sum_{j=1}^{N} n_j \quad (25)$$

To solve this problem, a two stage solution algorithm is proposed here. In the first stage, the main idea is to identify the central values of the unknown affine forms by first considering the energy hub operating at its nominal condition, which defines these central values. In this case, parameter uncertainties are not considered and thus the corresponding solution can be computed by solving the following scalar and deterministic optimization problem:

$$
\begin{aligned}
&\min_{\mathbf{x}_0} f_0 = f(\mathbf{x}_0) \\
&g_i(\mathbf{x}_0) = g_0^i = 0 \quad i = 1,..,n_g \quad (26) \\
&h_k(\mathbf{x}_0) = h_0^k \leq 0 \quad k = 1,..,n_h
\end{aligned}
$$

In the second stage, the effect of data uncertainty is considered, computing the partial deviations of the unknown affine forms by solving the following scalar and deterministic optimization problem:

$$
\min_{(x^i_{L_1},..x^i_{L_M},x^i_{cE_1},..,x^i_{cE_N},x^i_{Er_1},..,x^i_{Er_{Ng}})} rad(\hat{f}(x^i_{L_1},..x^i_{L_M},x^i_{cE_1},..,x^i_{cE_N},x^i_{Er_1},..,x^i_{Er_{Ng}}))
$$

$$
\begin{cases}
\sum_{j=1}^{M} g^i_{L_j} e_{L_j} + \sum_{j=1}^{N} g^i_{cE_j} e_{cE_j} + \sum_{j=1}^{N_g} g^i_{Er_j} e_{Er_j} \leq g^i_{n_x+1} - g^i_0 \\
-\sum_{j=1}^{M} g^i_{L_j} e_{L_j} + \sum_{j=1}^{N} g^i_{cE_j} e_{cE_j} + \sum_{j=1}^{N_g} g^i_{Er_j} e_{Er_j} \leq g^i_{n_x+1} + g^i_0
\end{cases} \quad i = 1,..,n_g \quad (27)
$$

$$
\sum_{j=1}^{M} h^k_{L_j} e_{L_j} + \sum_{j=1}^{N} h^k_{cE_j} e_{cE_j} + \sum_{j=1}^{N_g} h^k_{Er_j} e_{Er_j} \leq h^k_{n_x+1} - h^k_0 \quad k = 1,..,n_h
$$

The solution of these problems allows the analyst to identify the affine forms of the decision variables, which minimise the mean and the deviation of the objective function and satisfies the energy hub constraints for every operating point. The knowledge of these affine forms represents a strategic tool for defining effective operation scheduling strategies.

In details, they allow the analyst to assess an enclosure of the objective function range, which is guaranteed to contain its actual value. Moreover, the energy hub dispatching coefficients for the analysed time period can be defined according to the central value of the affine forms of the decision variables, namely:

$$\begin{cases} \mathbf{E} = \left[E_{1,0}, .., E_{N,0} \right]^T \\ E_{i,0} = \sum_{j=1}^{n_i} E^{j}_{i,0} \quad i = 1,..N \end{cases} \quad (28)$$

while the corresponding partial deviations can be used to correct the dispatching coefficients depending by the actual energy hub operation point. This can be obtained by computing the following discrepancy factors:

$$\begin{cases} \varepsilon^*_{L_j} = \dfrac{\bar{L}_j - L_{j,0}}{L_{j,1}} \quad j = 1,..,M \\ \varepsilon^*_{cE_j} = \dfrac{c\bar{E}_j - cE_{j,0}}{cE_{j,1}} \quad j = 1,..,N \\ \varepsilon^*_{Er_j} = \dfrac{\bar{E}r_j - Er_{j,0}}{Er_{j,1}} \quad j = 1,..,N_g \end{cases} \quad (29)$$

Where $[\bar{L}_1,..,\bar{L}_M, c\bar{E}_1,..,c\bar{E}_N, \bar{E}r_1,..,\bar{E}r_{N_g}]$ are the actual values of the uncertain input variables.

Then if the dispatch factors are fixed according to the following equations:

$$y^*_i = x^i_0 + \sum_{j=1}^{M} x^i_{Lj} \varepsilon^*_{L_j} + \sum_{j=1}^{N} x^i_{cEj} \varepsilon^*_{cE_j} + \sum_{j=1}^{Ng} x^i_{Erj} \varepsilon^*_{Er_j} \quad i = 1,..,n_Y \quad (30)$$

It follows that both the equality and inequality constraints are satisfied and the actual value of the cost function is guaranteed to be inside the computed range. These important features allow the analyst to effectively address simultaneously both the short/medium (i.e. one day ahead) and the real time operation-scheduling problem.

In particular, thanks to the Invariance Theorem of Affine Arithmetic, we can argue that the adoption of the AA based operators allows to compute an outer estimation of the cost and constraint functions bounds. In other words the bounds computed by AA are guarantee to include the real function domains. Anyway, the approximation errors induced by the application of non-affine operations may leads to an overestimation of the real bounds, leading to a loss of optimality in satisfying the Karush–Kuhn–Tucker (KKT) conditions, which roughly depends on the number of non-linear functions instances characterizing the optimization problem. This is a well note problem characterizing any range-based method for uncertain optimization, which is not expected to pose critical issues for the problem under study due to the adoption of a linearized model for describing the input/output energy hub equations.

Finally, it should be note that the computational costs of the proposed method sensibly depend on the number of noise symbols adopted for uncertainty representation, which could pose some computational difficulties for large-scale applications. To address this problem, advanced methods based on Principal Component Analysis for knowledge discovery from historical operation data-sets are currently under investigation by the first author of this paper. The main idea is to exploit the capacity of PCA in detecting potential relations among a set of operation data, which allows to describe the evolution of a large number of statistically correlated variables by a linear combination of a limited number of "primitive" variables. This feature is particularly useful in solving the problem formalized in (24), where the hypothesis of statistically independence of the uncertain parameters could involve the definition of a large number of noise symbols, which increases the complexities of the AA-based computations.

4.0 Simulation studies

This section discusses the application of the proposed methodology in the task of solving the optimal scheduling problem for the energy hub schematically depicted in fig.1. The considered hub architecture is based on the integration of a wind generator, a power transformer, a gas furnace and a combined heat and power unit (CHP). It processes two energy carriers (namely the electricity, $E_1 = E_{11} + Er_1$, and the natural gas, $E_2 = E_{21} + E_{22}$), and it supplies an electrical and a thermal load, denoted by L_1 and L_2 respectively. The corresponding coupling matrix is defined as follow:

$$\begin{bmatrix} L_1 \\ L_2 \end{bmatrix} = \begin{bmatrix} \eta_{11}^1 & \eta_{21}^2 & 0 \\ 0 & \eta_{22}^2 & \eta_{22}^3 \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{21} \\ E_{22} \end{bmatrix} + \begin{bmatrix} Er_1 \\ 0 \end{bmatrix} \quad (31)$$

Where the converter efficiencies are summarized in table I.

Table I: The Converter Efficiencies

Component	*Efficiency*
Power Transformer	$\eta_{11}^{1} = 0.98$
Combined Heat and Power Unit	$\eta_{21}^{2} = 0.35 \quad \eta_{22}^{2} = 0.405$
Gas Furnace	$\eta_{22}^{3} = 0.612$

Consequently the following vector describes the energy dispatching policy:

$$[E_{11}, E_{21}, E_{22}] \quad (32)$$

Subject to the following inequality constraints:

$$\begin{cases} 0 \le E_1 \le 20 \\ 0 \le E_2 \le 20 \\ 0 \le E_{11} \le 20 \\ 0 \le E_{22} \le 20 \\ 0 \le E_{21} \le 20 \end{cases} \quad (33)$$

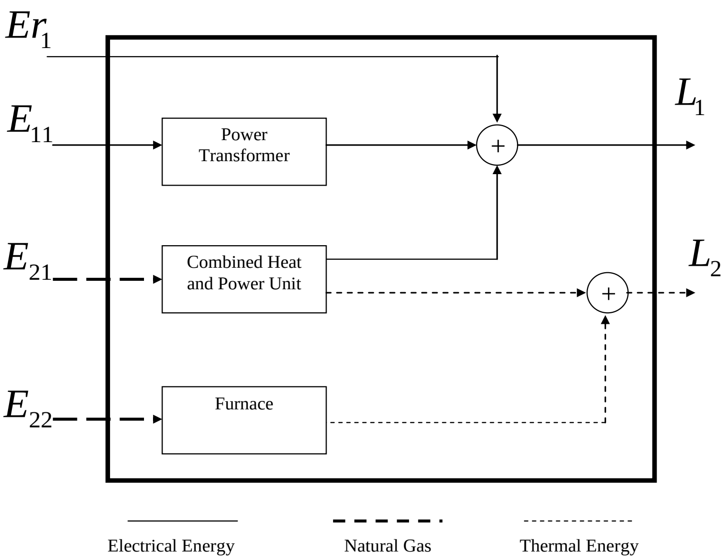


Fig.1: The Energy Hub architecture assumed in the simulation studies

As far as the data uncertainty are concerned, we consider the effect of electrical and thermal load forecasting uncertainty (assumed as $\pm 5\%$ and $\pm 2\%$ of the forecasted value respectively), wind energy forecasting uncertainty (assumed as $\pm 10\%$ of the forecasted value [30]) and electrical energy cost uncertainty (assumed as $\pm 10\%$ of the forecasted value [31]). The imbalance penalties have been modelled according to the following paradigm:

- If the energy imbalance is positive (namely the energy hub demands more energy respect to the scheduled value) then a cost increase of 30% has been assumed for quantifying the exceeding electrical energy cost and a cost increase of 10% has been assumed for quantifying the exceeding natural gas cost.
- If the imbalance is negative, then the user should pay the scheduled energy.

The first experiment deals with the solution of the optimal scheduling problem described by the following affine forms:

$$\begin{cases} \hat{L}_1 = 10.230 + 0.511\varepsilon_{L_1} = [9.719, 10.741]\,\text{MWh} \\ \hat{L}_2 = 11.640 + 0.233\varepsilon_{L_2} = [11.407, 11.873]\,\text{MWh} \\ \hat{E}r_1 = 1.055 + 0.105\varepsilon_{Er_1} = [0.950, 1.160]\,\text{MWh} \\ \text{c}\hat{\text{E}}_1 = 43.660 + 4.366\varepsilon_{cE_1} = [39.294, 48.026]\,\text{CAD/MWh} \end{cases} \quad (34)$$

Starting from these input data the proposed AA based solution methodology identifies the following solution:

$$\begin{cases} \hat{E}_{11} = 8.727 - 0.106\varepsilon_{Er_1} + 0.522\varepsilon_{L_1} = [8.099, 9.355] \\ \hat{E}_{21} = 1.774 \\ \hat{E}_{22} = 17.846 + 0.380\varepsilon_{L_2} = [17.466, 18.226] \end{cases} \quad (35)$$

Which allows us to define the following energy sourcing scheduling:

$$\begin{cases} E_{1,0} = 8.727 MWh \\ E_{2,0} = 17.746 MWh \end{cases} \quad (36)$$

The corresponding energy costs are guaranteed to lie in the interval $[894.62, 1021.9]$.

The solution described in (35) allows to define a real time strategy aimed at managing the deviation of the actual input data from the forecasted values. For example if the actual electrical load $\bar{L}_1 = 10.000\ MWh$ then the input energy flow to the first converter (namely the power transformer) should vary according to the following equation:

$$\begin{cases} \varepsilon_{L_1}^* = \dfrac{10.000-10.230}{0.511} = -450.098 \\ E_{11}^* = 8.727 + 0.522\varepsilon_{L_1}^* = 8492.049 MWh \end{cases} \quad (37)$$

To assess the robustness of the identified solution a Monte Carlo simulation aimed at randomly sampling the uncertain input data (34) has been implemented. For each data sample the dispatching coefficients (35) have been corrected according to (29) and the corresponding energy cost has been computed. The estimated bounds of the energy cost after 10000 trials are $[897.52, 1018.0]$. Analyzing these data it is worth observing as the proposed methodology allows us to compute a reliable enclosure of the cost function range.

As far as the computational requirements are concerned, the AA based solution strategy required about 2 seconds (on a 1.2 GHz CPU with 2 GB of RAM) to converge to a suitable solution for this case study. This is about 4% of the simulation time required by the stochastic programming method. In order to have a term of comparison for the performance evaluation of the proposed methodology, the optimal scheduling problem has been solved by a stochastic programming method based on Monte Carlo simulations. According to this approach, the bounds of the dispatching factors are inferred directly from repeatedly solving the deterministic problem (3) by randomly sampling the uncertain input data (34).

The analysis of the simulation results summarized in table II is based on two considerations:

1. by definition the Monte Carlo method, being sampling based technique, do not produce any spurious trajectories.
2. we make the hypothesis that the number of Monte Carlo trials is large enough to assume that the union of the uncertainty region described by this method is a very closed approximation of the correct problem solution.

Table II: Result Comparisons

Solution Methodology	*Energy Cost [CAD]*
AA	[894.62,1021.9]
Stochastic programming (1000 trials)	[896.05,1000.6]

Analyzing these data it is worth noting as the AA-based methodology gives fairly good approximations of the objective function bounds when compared to the benchmark intervals obtained with the Monte Carlo approach; this is mainly due to the intrinsic characteristic of AA that keeps track of correlations between the uncertain variables.

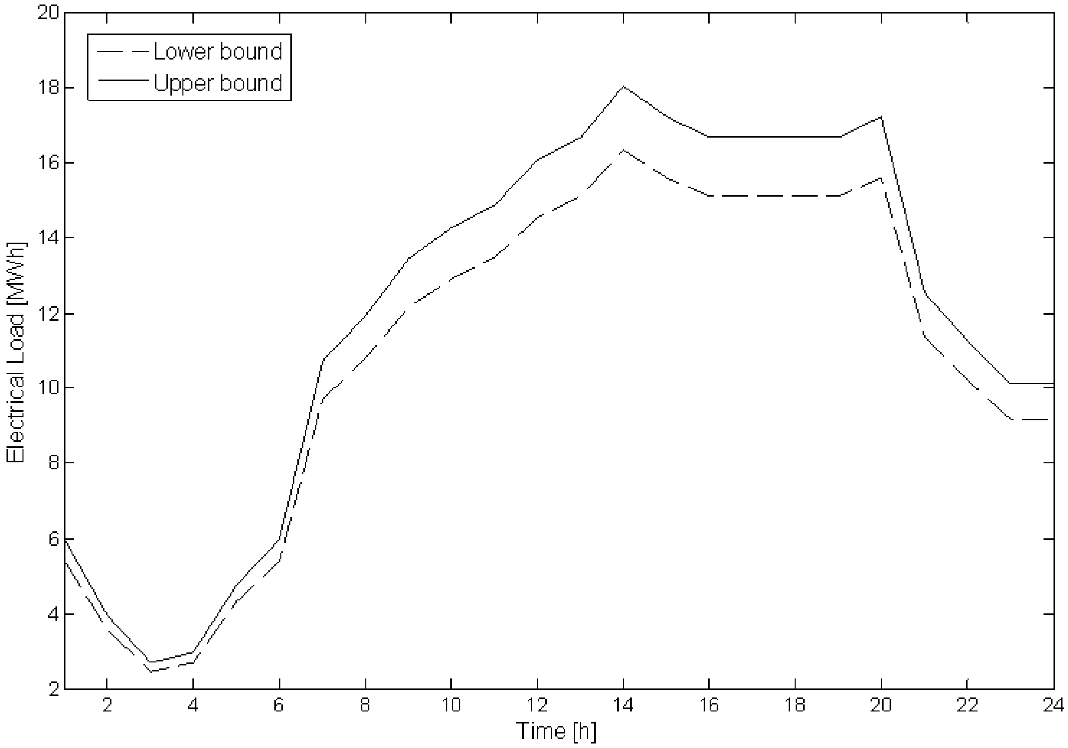
Notice also that the solution bounds are slightly conservative (of the order of 1%), which is due to the fact that AA yields "worst case" bounds, which take into account any uncertainties in the input data as well as all internal truncation and round off errors. This is to be expected, since the random, uniformly distributed variation of parameters assumed in the Monte Carlo approach tends to underestimate the worst-case variations. This can be considered an advantage of the proposed approach, since no assumptions regarding the probability distribution of the input uncertain data are required.

These benefits have been confirmed by further simulation studies aimed at solving the energy hub optimal scheduling problem for a 24h operating scenario characterized by the hourly ranges of the input data depicted in fig. 2[1]. The obtained results have been summarized in fig.3-5. In details in fig. 3 the central value and the upper/lower bounds of the hourly input energy flows are reported. The corresponding ranges of the hourly energy cost are depicted in fig.4. The same figure reports also the bounds computed by a Monte Carlo simulation with 10000 trials. In this context it is important to observe that this number has been determined by considering that after 10000 simulations we expected a substantial saturation in the assessment of the solution bounds, as it can be argued by analyzing the figure 5, which depicts the evolution of the upper bound of the cost function estimated by the Monte Carlo algorithm versus the number of trials, for a fixed hour.

The comparison of these two profiles confirm the effectiveness of the proposed methodology in solving the optimal energy hub scheduling problem also for highly variable load, generation and cost patterns.

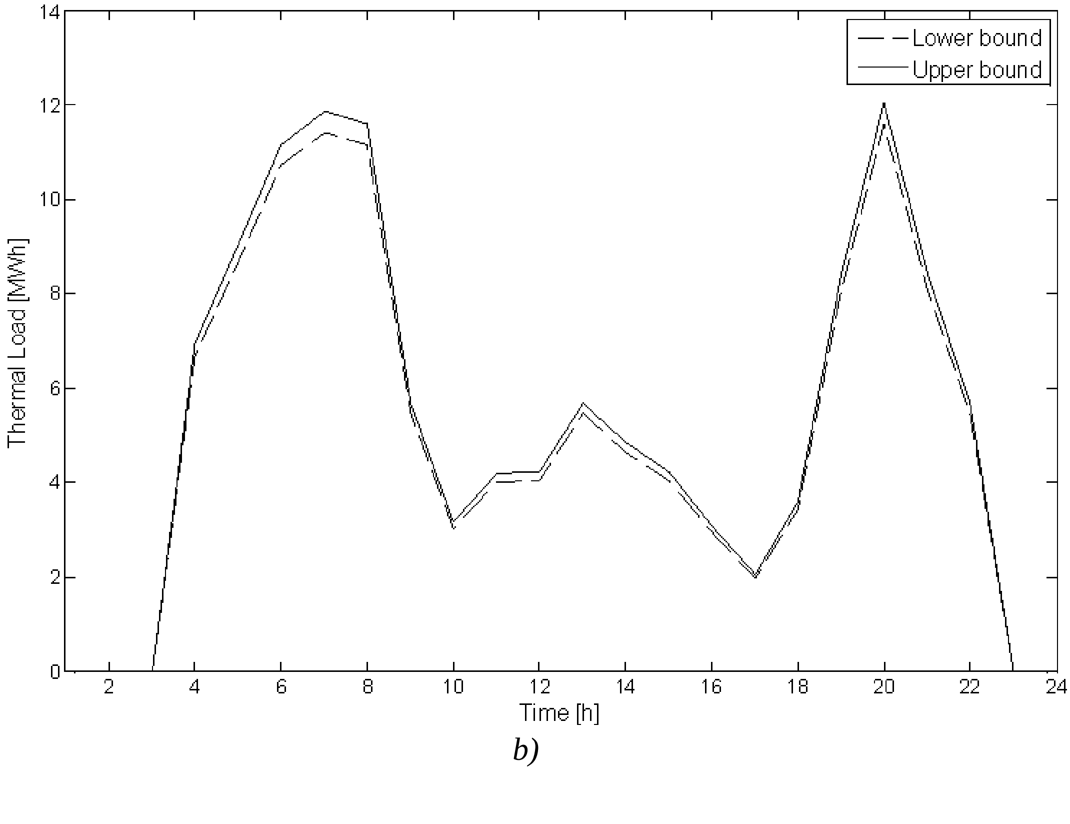
---

[1] The data assumed in the simulation studies have been extrapolated by processing the evolution of the energy prices in the Ontario market, and by assuming the energy hub profiles defined in [8,9].



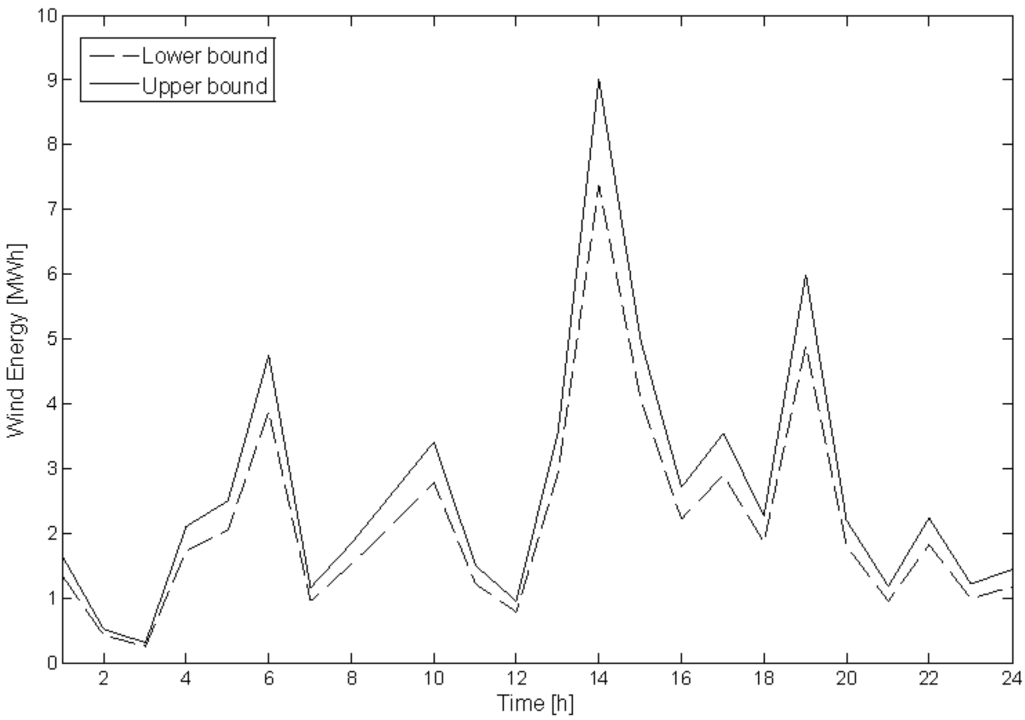
Lower bound
Upper bound
Electrical Load [MWh]
Time [h]

a)

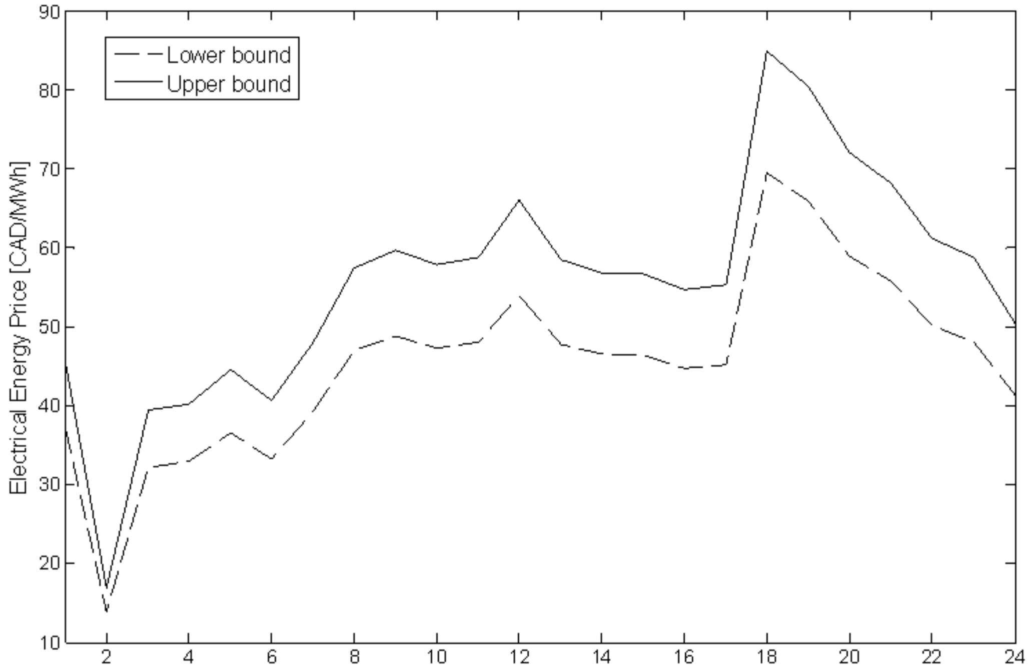


Lower bound
Upper bound
Thermal Load [MWh]
Time [h]

b)



c)



d)

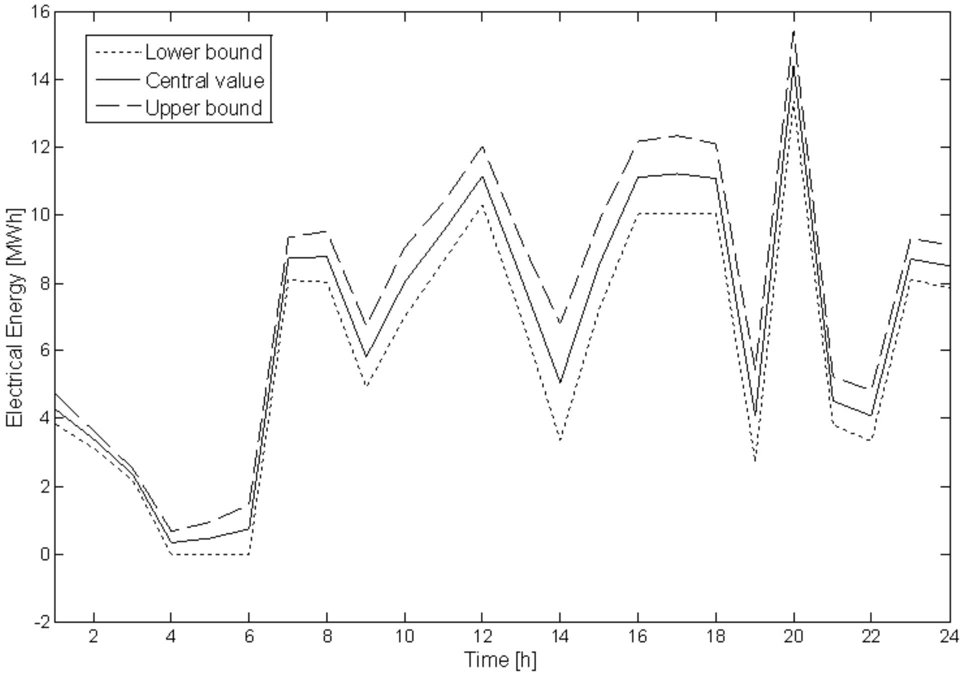
Fig. 2: Hourly input data bounds:

a) Electrical load

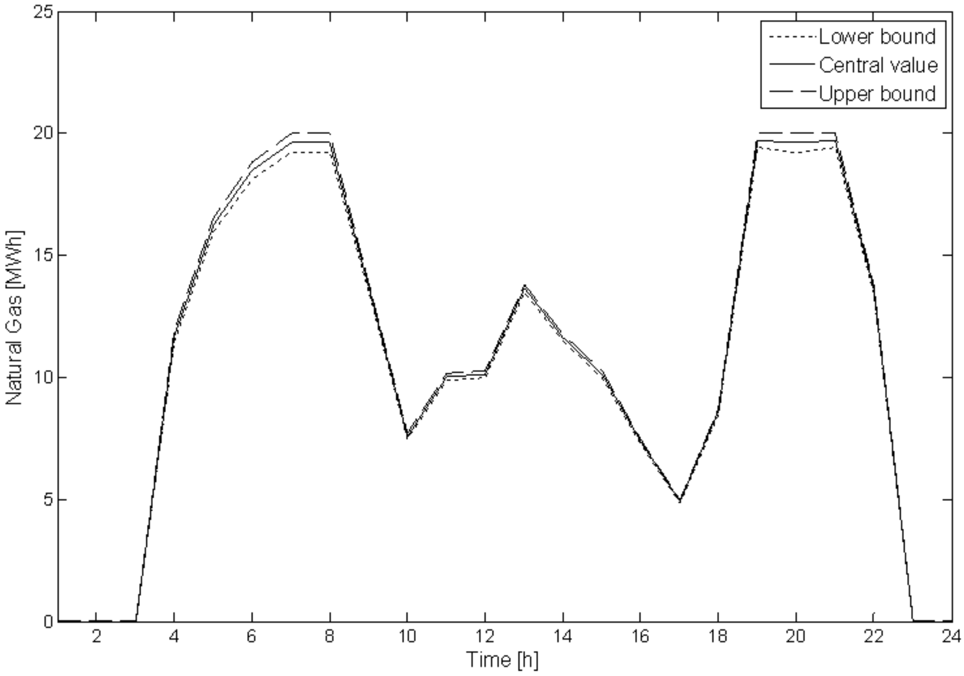
b) Thermal load

c) Wind energy

d) Electrical energy cost



a)



b)

Fig.3: Hourly bounds of the input energy flows:

a) Electrical energy

b) Natural gas

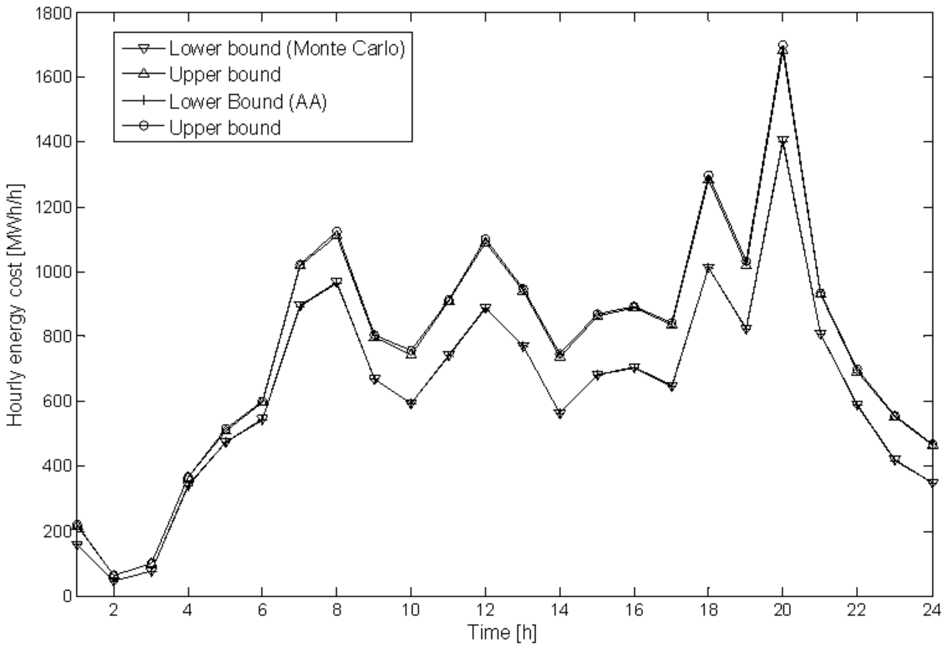


Fig. 4: Hourly bounds of the objective function

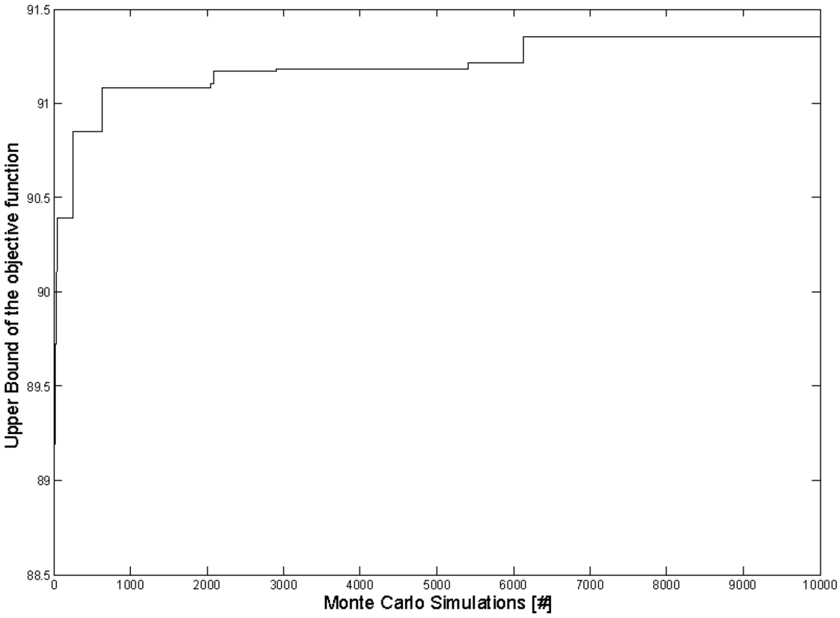


Fig. 5: *Upper bounds of the objective function estimated by the Monte Carlo based method in function of the number of trials for a fixed hour (e.g. 3 a.m.)*

5.0 Conclusive remarks

This paper proposed a new methodology for reliable energy hub operation scheduling in the presence of data uncertainty based on AA, allowing to better handle uncertainty compared to the traditional stochastic based solution approaches.

Based on the proposed new AA formalism, an enclosure of the objective function range, which is guaranteed to contain its actual value, was shown to be obtained by solving a constrained linear multiobjective programming problem. The solution of this problem by a two stage solution strategy allowed to effectively define reliable dispatching strategies aimed at simultaneously addressing both the short/medium and the real time energy hub operation scheduling problem.

The presented analyses and results demonstrate that the proposed AA-based approach is well suited for the assessment of uncertainty propagation in energy hub operation, and we expect that it can be effectively applied to study interconnected energy hub systems, independent of the types and levels of uncertainties in the input data.

The generalization of the proposed AA based optimization framework aimed at considering the effect of storage systems is currently under development by the authors.

References

[1] Martin Geidl and Göran Andersson, Optimal Power Flow of Multiple Energy Carriers, IEEE Transactions on Power Systems, 22/1 (2007), 145-155.

[2] T. Krause, G. Andersson, K. Fröhlich, A. Vaccaro, Multiple-Energy Carriers: Modeling of Production, Delivery, and Consumption, IEEE Proceedings, 99/1(2011), 15 - 27

[3] F. Kienzle, P. Ahcin, G. Andersson, Valuing Investments in Multi-Energy Conversion, Storage, and Demand-Side Management Systems under Uncertainty, IEEE Transactions on Sustainable Energy, 2/2, (2011), 194 – 202.

[4] Bozchalui, M.C., Hashmi, S.A., Hassen, H.,; Canizares, C.A., Bhattacharya, K., Optimal Operation of Residential Energy Hubs in Smart Grids, IEEE trans. on Smart Grid, Vol. 3(4) , pp. 1755 - 1766

[5] Ghasemi, A., Hojjat, M., Javidi, M.H., Introducing a new framework for management of future distribution networks using potentials of energy hubs, proc. of 2012 2nd Iranian Conference on Smart Grids (ICSG), 2012, Teheran, Iran.

[6] Schulze, M., Friedrich, L., Gautschi, M., Modeling and optimization of renewables: applying the Energy Hub approach, proc. of IEEE International Conference on Sustainable Energy Technologies, 2008, ICSET 2008, pp. 83 – 88

[7] Fabrizio E., Filippi M., Torbino M.,Operational optimization of actual energy systems by means of the energy hub theory, proc. of Building Simulation 2011, pp. 2793-2798, 14-16 November 2011, Sydney (Australia)

[8] A. Parisio, C. del Vecchio, G. Velotto, Robust Optimization of operations in energy hub, proc. of the 2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), pp. 4943 – 4948, 12-15 Dec. 2011, Orlando, FL

[9] A. Parisio, C. del Vecchio, A. Vaccaro, A Robust Optimization Approach to Energy Hub Management, International Journal of Electrical Power and Energy Systems, Elsevier, 42/1 (2012), 98–104

[10] F. L. Alvarado, Y. Hu, and R. Adapa, Uncertainty in power system modeling and computation, in Proc. IEEE Int. Conf. Systems, Man and Cybernetics, Oct. 1992, vol. 1, pp. 754–760.

[11] L. V. Barboza, G. P. Dimuro, and R. H. S. Reiser, Towards interval analysis of the load uncertainty in power electric systems, in Proc. 8th Int. Conf. Probabilistic Methods Applied to Power Systems, Ames, IA, Sep. 2004, Iowa State Univ.

[12] G., Bontempi, Simulating Continuous dynamical systems with uncertainty: the probability and the possibility approaches, In "Fuzzy Partial Differential Equations and Relational Equations", Editors: Masoud Nikravesh and Lotfi A. Zadeh, Series Studies in Fuzziness and Soft Computing, Physica-Verlag, Springer, 2003.

[13] B. R. Moore, “Methods and applications of interval analysis,” in SIAM Studies in Applied Mathematics. Philadelphia, PA: SIAM, 1975.

[14] L. H. de Figueiredo and J. Stolfi, Self-validated numerical methods and applications, in Proc. Brazilian Mathematics Colloq. Monograph, IMPA, Rio de Janeiro, Brazil, 1997.

[15] L. H. de Figueiredo and J. Stolfi, Affine arithmetic: Concepts and applications, Numer. Alg., vol. 37, no. 1–4, pp. 147–158, 2004.

[16] A.Vaccaro, C. Canizares, D.Villacci, An Affine Arithmetic based Methodology for Reliable Power Flow Analysis in the presence of Data Uncertainty, IEEE Trans. On Power Systems, Vol. 25(2), pp: 624 – 632, May 2010.

[17] J. Ninin, F. Messine, A Mixed Affine Reformulation Method for Global Optimization, Toulouse Global Optimization (TOGO 2010), p. 101-104, Toulouse, September 2010.

[18] E. de Rocquigny, N. Devictor, S. Tarantola, Uncertainty in Industrial Practice: A Guide to Quantitative Uncertainty Management, Ed. Wiley, June 2008

[19] P. Smets, Imperfect information: Imprecision and uncertainty, in Uncertainty Management in Information Systems, pp.225-254, Ed. Springer, 1997

[20] C. Boonchuay, K. Tomsovic, F. Li, and W. Ongsakul, Robust optimization-based dc optimal power flow for managing wind generation uncertainty, AIP Conference Proceedings, vol. 1499, no. 1, 2012.

[21] M. Mohammadi, A. Shayegani, and H. Adaminejad, A new approach of point estimate method for probabilistic load flow, International Journal of Electrical Power Energy Systems, vol. 51, no. 0, pp. 54 – 60, 2013.

[22] G. Shafer, A mathematical theory of evidence, vol. 1. Princeton university press Princeton, 1976.

[23] L. Hao, A. K. Tamang, Z. Weihua, and X. Shen, Stochastic information management in smart grid, IEEE Communications surveys Tutorials, vol. 16, no. 3, pp. 1746– 1770, 2014.

[24] P. Bijwe and G. V. Raju, Fuzzy distribution power flow for weakly meshed systems, IEEE Transactions on Power Systems, vol. 21, no. 4, pp. 1645–1652, 2006.

[25] R. Moore, Methods and applications of interval analysis, vol. 2. SIAM, 1979.

[26] L. V. Barboza, G. P. Dimuro, and R. H. Reiser, Towards interval analysis of the load uncertainty in power electric systems, in Proceedings of the International Conference on Probabilistic Methods Applied to Power Systems, pp. 538–544, 2004.

[27] M. Neher, From interval analysis to taylor models-an overview, Proceedings of the International Association for Mathematics and Computers in Simulation, 2005.

[28] Y. H. Wan, B. K. Parsons, Factors relevant to utility integration of intermittent renewable technologies, National Renewable Energy Laboratory, 1993.

[29] Soman, S.S., Zareipour, H., Malik, O., Mandal, P.,A review of wind power and wind speed forecasting methods with different time horizons, in proc. of North American Power Symposium (NAPS), 2010, Arlington, TX

[30] C. W. Lou, M. C. Dong, Modeling data uncertainty on electric load forecasting based on Type-2 fuzzy logic set theory, Engineering Applications of Artificial Intelligence, vol. 25, (8), December 2012, pp. 1567–1576

[31] Teoh, C.-C., Sheble, G.B, Applying Spot Rate Forecast Technique in Estimating Energy Price Uncertainties, proc. of International Conference on Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006, Stockholm, Sweden.