Titre: Title:	Consideration of marginal electricity in real-time minimization of distributed data centre emissions
Auteurs: Authors:	Thomas Dandres, Reza Farrahi Moghaddam, Kim Khoa Nguyen, Yves Lemieux, Réjean Samson et Mohamed Cheriet
Date:	2017
Référence: Citation:	Dandres, Thomas, Farrahi Moghaddam, Reza, Nguyen, Kim Khoa, Lemieux, Yves, Samson, Réjean et Cheriet, Mohamed (2017). Consideration of marginal electricity in real-time minimization of distributed data centre emissions*Journal of Cleaner Production*, *143*, p. 116-124. doi:10.1016/j.jclepro.2016.12.143

Consideration of marginal electricity in real-time minimization of distributed data centre emissions

Thomas Dandres[12], Reza Farrahi Moghaddam[2,3], Kim Khoa Nguyen[2], Yves Lemieux[3], Réjean Samson[1] and Mohamed Cheriet[2]

[1]: CIRAIG, Polytechnique Montréal, Montreal, Canada

[2]: Synchromedia, Ecole de technologie supérieure, Montreal, Canada

[3]: Ericsson Montreal, Montreal, Canada

Corresponding author: thomas.dandres@polymtl.ca

Abstract— **Among the innovative approaches to reduce the greenhouse gas (GHG) emissions of data centres during their use phase, cloud computing systems relying on data centres located in different regions appear promising. Cloud computing technology enables real-time load migration to a data centre in the region where the GHG emissions per *kWh* are the lowest. In this paper, we propose a novel approach to minimize GHG emissions cloud computing relying on distributed data centres. Unlike previous optimization approaches, our method considers the marginal GHG emissions caused by load migrations inside the electric grid instead of only considering the average emissions of the electric grid's prior load migrations. Results show that load migrations make it possible to minimize marginal GHG emissions of the cloud computing service. Comparison with the usual approach using average emission factors reveals its inability to truly minimize GHG emissions of distributed data centres. There is also a potential conflict between current GHG emissions accounting methods and marginal GHG emissions minimization. This conflict may prevent the minimization of GHG emissions in multi-regional systems such as cloud computing systems and other smart systems such as smart buildings and smart-grids. While techniques to model marginal electricity mixes need to be improved, it has become critical to reconcile the use of marginal and average emissions factors in minimization of and accounting for GHG emissions.**

Key words: **Distributed data centre optimization; GHG emissions; real-time electricity generation; marginal electricity.**

1. Introduction

Information and communication technologies (ICTs) have grown exponentially in the last decades and this rapid growth is expected to continue (Gantz & Reinsel, 2012). However, manufacture and use of ICTs are associated with large electricity and resources consumption (Lannoo et al., 2013). In 2006, ICTs were found to contribute to 2% of global anthropogenic greenhouse gas (GHG) emissions, which were equivalent to the emissions of the aviation industry for that year (The Climate Group, 2008). Because data centres are one of the three major sinks of electricity among ICT infrastructures, they also significantly contribute to ICT GHG emissions (Lannoo et al., 2013). Therefore, significant effort has been invested to curb data centre electricity demand, improve their efficiency and reduce their environmental footprint (Beloglazov et al., 2012; Doyle et al., 2013; Van Heddeghem et al., 2012).

Among the innovative approaches to reduce data centre use-phase GHG emissions is overall load management across distributed data centres (Amokrane et al., 2013; Krioukov et al., 2011; Mandal et al., 2013). In this approach, data centres are located in several regions and connected to the regional electrical grid. Load management is used to vary the power demand of the data centres in real time to maximize power consumption in regions where the GHG emissions per *kWh* are the lowest. Indeed, electricity is generated to instantly meet the regional power demand, which changes continuously during the day, depending on consumer needs. Therefore, the regional mix of power plants changes constantly, as does the related GHG emissions factor per *kWh*. Thus, there is an opportunity for distributed data centers to minimize their GHG emissions in real-time by identifying the least emitting sources of electricity.

1.1. Problem statement

According to our knowledge, in this innovative approach, the choice of region to which the load is migrated to is related to real-time electricity generation data (in the best cases). Concretely, the electrical grid mixes several regions are checked regularly and then the load is balanced between regions where the GHG emissions are lower at a given time. This means that the changes in regional power demand caused by the load balancing are not taken into account in this approach, since the load balancing is made after the grid mix check. Consequently, the regional grid mix change (and its consequences on regional GHG emissions) directly caused by load management is also ignored. In other words, the current load balancing approach would not capture an increase in regional coal power generation caused by a rise in the load processed by a data centre in that region. Thus, there is uncertainty regarding the real GHG emissions reductions achieved (if any) with the optimization of data centre networks when using the current load balancing approach. Therefore, a method adapted to the dynamic electricity context is needed to instantly identify power plant types affected by load balancing and minimize the GHG emissions of distributed data centres.

1.2. Objectives

Thus, the objectives of this study are to:

- Highlight the need to consider marginal sources of electricity when optimizing distributed data centres using load balancing,
- Present approaches to identify marginal sources of electricity,
- Discuss issues related to multiregional optimization based on marginal electricity management and GHG emissions accounting.

A case study has been built to illustrate these objectives. In this case study the GHG emissions of a cloud computing service are minimized using load balancing between distributed data centres located in two different Canadian provinces.

2. Method

This section begins with a description of the case study used to illustrate the role of marginal electricity in a Canadian data centre network. Then, the method used to identify the marginal sources of electricity is presented. Finally, several scenarios are defined to evaluate the method's performance.

2.1. The Green Sustainable Telco Cloud

The case study involves the Green Sustainable Telco Cloud (GSTC) that is defined as a cloud computing service based on an efficient, optimized and environmentally-friendly distributed data centre network. Several optimization criteria, such as service quality, GHG emissions and operating costs, are considered. However, this paper focuses only on the environmental criteria to fully illustrate the role of the marginal sources of electricity in GHG emissions. The GSTC is currently under development, and the case study presented here is therefore based on a scenario rather than real results. It was deliberately simplified and is more conceptual than practical. Thus, it does not present accurate and complete GHG emissions results.

In this case study, two virtualized data centres are located in the Canadian provinces of Ontario and Alberta and form a cloud computing system that provides online services. These provinces were chosen because detailed, historic and real-time electricity generation data are available from public sources (AESO, 2013; IESO, 2013). It is assumed that the two data centres are similar and connected to the regional grids and that only one handles the cloud computing service at a time. In addition, the cloud power demand varies over time on a daily basis, depending on user requests, as presented in Fig. 1. It is also assumed that data transmission by all users towards the data centres consumes the same amount of electricity, regardless of the location of the data centre hosting the cloud.

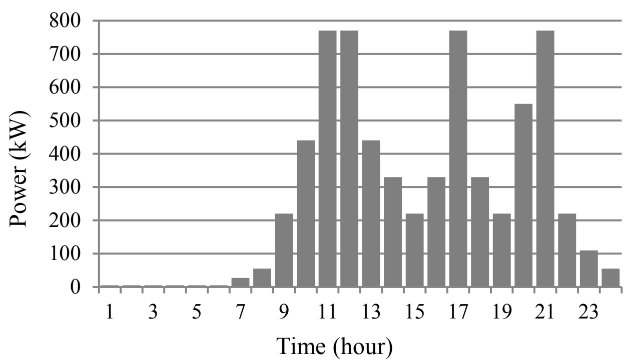

Fig. 1. Hourly cloud power requirement

It is considered that one load migration between the two data centres could be made every hour. Assuming that the GSTC provides an online service that host a negligible amount of user data, the load migrations should not cause significant additional data traffic. Consequently, the electricity consumption due to such data traffic has been neglected. Concretely, the GSTC service could consist of online file processing, such as picture or video editing or mathematical computing.

2.1.1. Definition of marginal electricity

Electrical grid networks are balanced between producers and consumers. In practice, the electrical grid operator must keep the power generation capacity close to the power demand. Since the power demand changes over time due to consumer behaviours, producers must constantly adapt their power generation capacity. Adapting power generation capacity in real time depends on various criteria such as the flexibility of the power generation technologies, the power plants' operating costs, the electrical grid constraints and the electrical grid operator schedules. By definition, a power plant that adapts its power generation capacity in response to a change in power demand is a marginal source of electricity. The marginal electricity is the electricity generated by all the marginal sources. Several observations in the electricity sector show that, for a very small variation in power demand, there is only one marginal technology (Nielsen et al., 2011; Sensfuß et al., 2008; Tveten et al., 2013). However, this technology may change over time, and more than one technology may be affected when the demand variation increases (Mathiesen et al., 2009; Rogers et al., 2013b). Moreover, since the marginal sources of electricity are selected among a pool of potential marginal sources of electricity, there is uncertainty in the prediction of the marginal sources of electricity. To overcome this problem, past data were used in this study to retrospectively identify the marginal sources of electricity. These data are retrieved from the Canadian electric utilities (AESO, 2013; IESO, 2013).

2.1.2. Assumption on GSTC power demand

To simplify electricity generation modeling, the cloud power demand in Ontario and Alberta is considered to be entirely met by marginal sources of electricity. A different and perhaps more realistic approach would be to split the cloud power demand between the static and varying power demands. Therefore, the static power demand would be met by non-marginal electricity and the varying power demand by marginal electricity. However, this assumption would complicate computing without providing relevant information to the case study.

2.1.3. Identification of marginal sources of electricity

The data collected from the Canadian utilities quantify the amount of electricity generated per power generation technology and per hour in Alberta and Ontario over 2011-2013. Using these data, the variations in generation of each technology was monitored for each hour. The result is an increase or a decrease in power generation per technology, per hour and per Canadian province. It constitutes the hourly marginal sources of electricity. An example of the calculation is provided in the supplementary material. By building a marginal power mix made of several technologies, this approach is in agreement with Mathiesen et al. (2009), who recommend not considering only one marginal technology. As presented in the supplementary material, the marginal technologies' decreasing or raising their capacity is expected to contribute in the same manner to the rises in power demand.

One simplification in the identification of marginal sources of electricity is that electricity imports were not considered. The marginal electricity in a region may come from abroad, since electricity providers can import electricity to overcome peak power demands. Indeed, it may be cheaper to import electricity at a high price than to build a power plant that will be used for only a few hours per year. As presented in supplementary material, electricity imports in Ontario and Alberta are low in comparison to total power generation. Moreover, imported electricity is expected to come from hydro dams with low GHG emissions per *kWh*. Thus, these imports should not change significantly the results of this study. Nevertheless, to be very rigorous, the marginal sources of electricity in interconnected regions should be identified and included in the regional power mix. However, implementing the electricity imports in the calculation would have made it unnecessarily complex without providing significant added value to this study.

2.1.4. Calculating the marginal GHG emissions factors

Once the marginal sources of electricity are identified, the hourly GHG emissions are computed for each region according to the amount of electricity provided by the marginal electricity mix and on the life-cycle GHG emissions factor of each technology. The life-cycle emissions factors were taken from the ecoinvent database (version 3.1) using Simapro (version 8.1) and IMPACT2002+ (2.20) for 1 *kWh* of electricity. Life-cycle emissions factors account for power plant construction as well as other life-cycle steps in electricity generation such as fuel extraction (Şengül et al., 2016). Thus life-cycle emissions factors are not restricted to energy extraction from fossil/fissile fuels or the transformation of renewable energy into electricity. The ecoinvent processes used to model the different technologies are presented in Table 1. The impacts per *kWh* of electricity transport and distribution were included on the basis of medium and low voltage electricity processes available in ecoinvent for Ontario and Alberta. Hourly marginal GHG emissions were then divided by the total amount of marginal electricity generated during the hour to obtain the hourly marginal GHG emissions factor per *kWh*. An example of the calculation is provided in the supplementary material.

Table 1
Modelling of electricity generation technologies in ecoinvent

Energy source	ecoinvent process
Biomass	Electricity, high voltage {CA-AB/ON}| heat and power co-generation, wood chips, 6400kW thermal, with extensive emission control | Alloc Def, U
Coal	Electricity, high voltage {CA-AB/ON}| electricity production, hard coal | Alloc Def, U
Hydro	Electricity, high voltage {CA-AB/ON}| electricity production, hydro, run-of-river | Alloc Def, U Electricity, high voltage {CA-AB/ON}| electricity production, hydro, reservoir, alpine region | Alloc Def, U
Natural gas	Electricity, high voltage {CA-AB/ON}| electricity production, natural gas, at conventional power plant | Alloc Def, U Electricity, high voltage {CA-AB/ON}| electricity production, natural gas, combined cycle power plant | Alloc Def, U
Nuclear	Electricity, high voltage {CA-ON}| electricity production, nuclear, pressure water reactor | Alloc Def, U
Oil	Electricity, high voltage {CA-AB/ON}| electricity production, oil | Alloc Def, U
Wind	Electricity, high voltage {CA-AB/ON}| electricity production,

wind, <1MW turbine, onshore | Alloc Def, U
Electricity, high voltage {CA-AB/ON}| electricity production, wind, 1-3MW turbine, onshore | Alloc Def, U
Electricity, high voltage {CA-AB/ON}| electricity production, wind, >3MW turbine, onshore | Alloc Def, U

2.1.5. Minimization of real-time GHG emissions

In order to minimize the real-time GHG emissions of the cloud computing service, the load was processed in the data centre located in the region where the hourly marginal GHG emissions factor per *kWh* was the lowest. Then, the GHG emissions related to the GSTC service were computed based on the cloud electricity consumption and the relevant hourly regional marginal emission factor. An example of the calculation is provided in the supplementary material.

2.2. Comparison with no-cloud computing scenarios

To assess the benefits of load balancing in the GSTC context, two scenarios without cloud computing systems were modelled. In these scenarios, it is presumed that the online service is hosted by only one data centre. The data centre is located in Ontario in one scenario and in Alberta in the other. Data centre GHG emissions for each scenario were calculated on the basis of the cloud computing electricity consumption (same as Fig.1) and regional GHG marginal emission factors computed previously. Then, the GHG emissions of the three scenarios (Ontario and Alberta, only Ontario and only Alberta) were compared.

2.3. Comparison with the conventional approach

A comparison with the conventional approach (using average emissions factors without considering marginal electricity) was conducted to highlight the differences between the optimization achieved when considering marginal electricity. To this end, the average GHG emission factors were computed for Ontario and Alberta for every hour over 2011-2013 based on the global electricity mix (instead of the marginal sources of electricity) following the method presented in section 2.1.4. Then, these average GHG emission factors were compared for each hour to minimize the GHG emissions of the cloud computing service following the method presented in section 2.1.5. Finally, the GHG emissions of the data centre in the two scenarios without cloud computing systems were computed using the relevant average GHG emissions factors and the cloud electricity consumption following the method presented in section 2.1.5. Examples of the calculations of the average electric mix, average emissions factor and minimization of average GHG emissions are provided in supplementary material.

3. Results

3.1. Minimization of real-time marginal GHG emissions

GHG emissions per marginal *kWh* and per context (cloud computing and non-cloud computing scenarios) are presented in Table 2. The marginal GHG emissions factors are in agreement with those of Farhat et al.(2010) computed for each Canadian province. Nevertheless, some differences are observed because Farhat et al.'s study was conducted using 2004-2006 data, while we used 2011-2013 data. Especially, the coal phase-out policy in Ontario can be observed in the 2011-2013 period. Thus, our marginal emission factors are inserted between the two scenarios (with and without coal) developed by Farhat et al. As illustrated, load migrations help in minimizing the cloud computing service GHG emissions. This reduction in GHG emissions is possible because marginal GHG emission maximums do not occur at the same time in Alberta and Ontario (especially since the provinces are in different time zones). By avoiding the marginal emission maximums of both regions, the mean of marginal GHG emissions per *kWh* decreases in the cloud computing scenario as compared to no-cloud computing. Concretely, the load migrations reduce GHG emissions by 9% as compared to the scenario in which a single data centre is located in Ontario and by 44% when a single data centre is located in Alberta. These results corroborate that there is a high potential of GHG emissions to be avoided when load migrations are managed properly, even when marginal electricity is considered.

Table 2
GHG emissions per marginal *kWh* and per cloud computing scenario

Marginal GHG emissions (*kg CO_2 eq.*)	Optimization	Ontario	Alberta
per marginal *kWh* (2011-2013):			
Mean	0.317	0.344	0.618
Standard deviation	0.165	0.185	0.211
per data centre:			
2011-2013	2,441,680	2,677,536	4,326,700
Cloud computing	Yes	No	No

The difference in GHG emissions per marginal *kWh* between Ontario and Alberta is explained by the difference in marginal electric mixes in these regions. While marginal electricity is usually generated from natural gas and hydro (low GHG emissions per *kWh*) in Ontario, Alberta generally relies more on natural gas and coal (high GHG emissions per *kWh*). Therefore, the mean GHG emissions per marginal *kWh* in Alberta are higher than in Ontario. Consequently, the GSTC is hosted more often in Ontario (82% of the time) than in Alberta. These results also highlight the need to check the regional marginal sources of electricity when deploying a system with time-dependent power demand.

3.2. *Comparison with the conventional approach*

GHG emissions per average *kWh* and per situation (cloud computing and no-cloud computing scenarios) are presented in Table 3. According to the conventional approach, the minimization of the GHG emissions of the cloud computing service results in processing the load in the data centre located in Ontario at all times (*Optimization* and *Ontario* columns are identical in Table 3). This result was expected since nuclear power (low GHG emissions per *kWh*) is the main source of electricity in Ontario versus coal (high GHG emissions per *kWh*) in Alberta. The maximum GHG emissions per average *kWh* in Ontario are therefore always lower than the minimum GHG emissions per average *kWh* in Alberta. Thus, the conventional approach never recommends choosing the data centre in Alberta to process the load. This is an important finding: the conventional approach does not always yield a fully optimized solution (regardless of the fact that the conventional approach may not accurately model the GHG emissions).

Table 3
GHG emissions per average *kWh* and per cloud computing situation

Average GHG emissions (*kg CO_2 eq.*)	Optimization	Ontario	Alberta
per average *kWh* (2011-2013):			
Minimum	0.036	0.036	0.667
Mean	0. 125	0.125	0.821
Maximum	0. 298	0.298	0.930
Standard deviation	0. 048	0.048	0.032
per data centre:			
2011-2013	1,020,138	1,020,138	5,982,113
Cloud computing	Yes	No	No

Regarding the emissions presented in Tables 2 and 3, GHG emissions per marginal *kWh* in Ontario are higher compared to GHG emissions per average *kWh* in Ontario. The main reason is that nuclear power contributes very little to marginal electricity generation but represents about half of the average power generation. Since nuclear power emits very few GHG emissions per *kWh*, the small contribution of nuclear power to the Ontario marginal power mix makes the marginal GHG emissions factor greater than the average one. In Alberta, the marginal GHG emissions factor is lower than the average one because many Alberta power plants burn coal continuously without considering the province's power demand variations. These power plants are therefore excluded from Alberta's marginal power mix. Since coal power plants release significant amounts of GHG per *kWh*, the marginal power mix in Alberta emits fewer GHG emissions per *kWh* than the average power mix.

It should be noted that the emissions per data centre in Tables 2 and 3 do not reflect the same reality. Emissions from marginal electricity (Table 2) represent emissions due to change in regional power demand while emissions from average electricity (Table 3) represent the emissions due to regional electricity consumption. The purpose of considering marginal electricity is to minimize the GSTC's emissions but not to account for them. Average GHG emissions factors are commonly used to account for GHG emissions. However, as discussed in the next section, the difference between the average and marginal emission factors has challenging implications when accounting for marginal emissions avoided by marginal electricity management.

4. Discussion

4.1. *The discussion begins with the description of the problematic of real-time minimization and assessment of emissions and then expands to broader topics such as marginal electricity mix prediction and other criteria related to the cloud computing optimization problem, and then concludes with a presentation of the global applicability of the method.Minimizing and accounting for the GHG emissions of multiregional systems*

Minimizing and accounting for the GHG emissions of a system managing marginal electricity in several regions such as the GSTC might be conflicting. Indeed, we recommend the use of marginal emission factors to minimize emissions, but emissions accounting usually implies the use of average emissions factors. The source of the problem is that these emissions factors may be different because they rely on different electricity mixes. Thus, minimizing marginal emissions from distributed data centres located in several regions may result in accounting for more average emissions. This case is actually observed in our study. Minimization of marginal GHG emissions results in 2,441,680 *kg CO_2 eq*. (see Table 2). Using the regional average GHG emissions factors to account for the emissions of the optimized cloud would result in 2,037,518 *kg CO_2 eq*. The problem is that these emissions are higher than 1,020,138 *kg CO_2 eq*. (see Table 3) corresponding to the case where the server load is processed all the time in Ontario. Despite some marginal GHG emissions can be avoided by load migrations, the GHG emissions accounting method based on average emission factors prevent load balancing between data centres located in Ontario and Alberta. Concretely, this means that accounting GHG emissions with average emission factors does not provide a motivation to reduce marginal GHG emissions and can potentially prevent such reduction. Especially, in the context of a carbon tax computed with average emissions factors, the minimization of marginal GHG emissions would double the amount of the carbon tax in this study. Thus, there is a risk that the current GHG emissions accounting method causes more emissions than expected. Indeed, it encourages electric consumption in regions where average GHG emission factors are low while underestimating marginal emissions caused by power demand variations. Beyond distributed data centres, this problem also concerns smart-grid applications that aimed to minimize emissions in real time. Especially, some authors already recommend using real-time marginal emission factors to evaluate and minimize GHG emissions of smart buildings (Roux et al., 2016), electric vehicles (McCarthy & Yang, 2010; Thomas, 2012) and users adapting their power demand (Zheng et al., 2015). Indeed, a low average emission factor at a given time in a region does not guarantee that the marginal sources of electricity will also have low emissions at that time. Therefore, there is a critical need to develop a new approach to account for GHG emissions of smart systems that minimize their marginal emissions in real-time. Yang (2013) lists and discusses the merits and drawbacks of different allocation methods of electricity generation emissions in the context of electric vehicles. He concludes there is no ideal allocation method that perfectly meets the need for accuracy, simplicity/transparency or consistency. A possible solution could be a hybrid method in which the allocation of the marginal and non-marginal emissions between all electricity consumers would depend on the steady and fluctuating parts of their power demands.

4.2. Prediction of marginal sources of electricity

Predicting marginal sources of electricity is required to manage marginal electricity. However, this task is difficult because of the complexity of electric networks where the power generation may vary greatly in time and space (Olkkonen & Syri, 2016). In an ideal situation, the regional electrical utilities would send to the cloud manager real-time information on the short-term marginal sources of electricity identified by their physical models (Raichur et al., 2016).

Until such relationships between utilities and cloud managers are established, however, it is necessary to predict the origin of the marginal electricity. While it could be difficult to predict exactly which power plant will adapt its power generation capacity in response to a change in the power demand, it may be sufficient to know which technology is affected by the power demand change, since the GHG emissions factor is mostly determined by the power plant's technology. The following sections discuss the constraints imposed on marginal electricity generation and the different approaches that may be used to identify marginal sources of electricity.

4.2.1. Factors affecting marginal electricity generation

The choice of marginal sources of electricity by the utilities is conditioned by at least two factors: the flexibility of the power generation capacity and the price of electricity on the short-term market. Ideally, prediction methods should take into account both of these factors.

The flexibility of the power generation capacity: For technological reasons, not all power plants can easily adapt their power generation capacity in real time. For instance, nuclear technology is not flexible and can hardly follow the hourly changes in power demand. Therefore, nuclear technology is used to supply base power demand and generates as much electricity as possible, regardless of the demand. Hydroelectric technology is more flexible and can quickly adapt its capacity. Thus, hydroelectricity is a suitable technology to manage rapid changes in power demand.

The price of electricity on the short-term market: In the context of an open market, the price of electricity in the short term (including operation, transmission and distribution prices) is a determining factor that affects the choice of the marginal source of electricity used to provide additional power (Nielsen et al., 2011; Sensfuß et al., 2008; Tveten et al., 2013). Indeed, the cheapest technologies are often selected first if their generating capacity can be adapted to the power demand. Since regions are usually interconnected, a marginal source of electricity can be located far away from the consumer causing a change in the power demand. This is especially true when a region faces a high power demand that surpasses the installed regional power generation capacity.

4.2.2. Methods to identify marginal technologies

There are several approaches to pinpoint short-term marginal electricity generation technologies: the economic modeling, historical data, market data and literature-based identification. However, these approaches depend highly on the availability of data. Indeed, electricity suppliers do not always share sufficiently accurate operation data. Therefore, the identification of the marginal electricity generation technologies may be more or less uncertain, depending on the quality of the data obtained. What is important is that the data reflect the reality of the real-time and short-term behaviours of electricity generation. Possible approaches to identify marginal technologies are summarized below.

Economic approach: This approach aims to calculate the price of marginal electricity for different technologies and identify the marginal technology on the basis of electricity price (Amor et al., 2011; Kopsakangas-Savolainen et al.). Ideally, data on the efficiency of each power plant in the considered region are required. However, technology efficiency may be sufficient. Then, the real-time stock market prices of fuels can be used to compute the price of marginal electricity. The flaw of this approach is that fuel stock market prices vary considerably, thus affecting the choice of marginal technology. Moreover, the installed generation capacity of each technology must be known in order to consider the availability of each marginal technologies in a region. Finally, the cheapest power generation technology maybe not the marginal one at a given time due to technical constrains on the electrical grid (Hawkes, 2010).

Historical approach: In this approach, the marginal technologies are identified according to an analysis of historical electricity generation data (Farhat & Ugursal, 2010; Siler-Evans et al., 2012). Hourly electricity generation data must therefore be available for a period of at least one year in order to take into account seasonal effects. Marginal technologies can then be identified for each hour in each season or month. The major drawback of this approach is that the past power mix may be not representative of the current one. For instance, new power plants and transmission lines (including interconnections with other networks) are not reflected in historical data.

Market approach: This approach identifies marginal technologies based on electricity market information. Electricity generation is scheduled in advance according to power demand predictions (usually the day before). Then, the predictions are corrected (a few hours before) and power generation is adjusted in real time. In liberalized electricity markets, there is generally a call for tenders among electricity providers. In theory, the identity of the winning bidders and the technologies schedule make it possible to determine the marginal technologies. However, it may be problematic to obtain such information in real-time because of confidentiality terms. Another type of market data that is commonly publicly shared in real-time is the marginal price of electricity. Therefore, a method has been developed to extrapolate the short-term marginal emissions from the real-time marginal electricity prices (Rogers et al., 2013a). A flaw of this approach is that it does not determine marginal technologies when the demand decreases.

Literature review approach: It is possible to rely on existing studies to determine the main marginal technology. Of a region. Nevertheless, while a literature review may help quickly identify marginal technologies, available studies are not always very recent and may lead to erroneous marginal technology determination. Indeed, like historical data, old studies may not reflect the current reality of an electrical network. Moreover, Mathiesen et al. (2009) report that a single marginal technology is often considered in some electricity studies despite the existence of a marginal mix of technologies. Because of the high variability of GHG emissions between power generation technologies, this simplification results in significant uncertainty on emissions in these studies.

Implementation of predictive methods

In recent projects, we implemented the historical approach to predict marginal sources of electricity in the context of Ontario/Alberta (Dandres et al., 2014) and Québec (Maurice, 2015). A high variability among the technologies used to generate marginal electricity was observed. But despite the uncertainty of the predictions, it was still possible to achieve some minimization of the GHG emissions. However, more research is needed to improve the predictive models developed as part of these projects. To that end, big data analytics, as used in other areas (Li et al., 2016), appears promising to improve our models.

4.3. Minimization of GHG emissions: perspective of the data centre owner

Suppose a situation where the GSTC would be hosted by private data centers with different owners. Assuming that each data centre owner wants to use his facility as much as possible to generate financial profits, they may be tempted to serve other consumers when the GSTC load is not processed in their data centres. In this situation, the environmental benefits of load balancing would be reduced if not cancelled. To overcome this problem, a financial incentive mechanism, like the one of Gaz Métro (Société en commandite Gaz Métro et al., 2007) (the main natural gas provider in Québec) to compensate for sales reductions when its consumers implement energy efficiency measures, may be considered. However, reaching this type of agreement could become complex when several governments of different regions are involved. Also, the possibility of generalizing the marginal approach of this paper to a profit-aware optimization (Farrahi Moghaddam, 2014) will be considered in the future.

4.4. Non-GHG emissions and environmental life-cycle impacts

Beyond GHG emissions, there are other impactful substances that may be emitted during electricity generation: fine particles that aggravates respiratory and cardiovascular diseases, aromatic hydrocarbons that have a carcinogenic effect on human health, sulphur and nitrogen dioxides that cause acid rain, etc. (Frischknecht & Rebitzer, 2005; Jolliet et al., 2003). Also, another optimization criteria such as the sustainability power density defined by Buceti (2014) would probably lead to different choices and consequently to different emissions. Distributed data centres' optimization based only on GHG emissions are therefore not expected to minimize damage to human health or ecosystems.

Moreover, the data centre's building and dismantling phases are also sources of environmental impacts (Meza et al., 2010; Whitehead et al., 2012). Thus, there is an environmental burden attributed to each data centre, and the environmental benefits obtained through load balancing must at least be equal to this burden if a data centre is built for that purpose. Life-cycle assessment (LCA) (ISO standards 14040-44) may be used to evaluate the environmental impacts of ICT equipment (Bonvoisin et al., 2012; Bonvoisin et al., 2014; Deng et al., 2011) and the environmental benefits of ICT virtualization (Mirabella et al., 2013). LCA makes it possible to model the potential environmental impacts of products in line with the emissions and resource consumption that occur at each stage in a product's life-cycle. The emissions and resource consumption are then converted into impacts for several environmental indicators using environmental impact assessment models. Moreover, LCA enables the comparison between different options, such as videoconferencing or face-to-face meetings (Borggren et al., 2013), and the design improvement of ICT devices by choosing materials friendlier to the environment (Meyer & Katz). However, using LCA to study ICTs presents certain challenges, such as collecting detailed data on manufacturing component and waste management or modelling the rebound and substitution effects of ICT solutions (Arushanyan et al., 2014; Farrahi Moghaddam et al., 2014; Mirabella et al., 2013).

4.5. Limits and future work

In this study, several simplifications were made and must be addressed in future work:

- Electricity imports were excluded from the study but should be considered when evaluating marginal electricity. To this end, the sources that export electricity to the studied regions must be specified (moreover, should all sources abroad be considered or only marginal sources?). Ecological footprints and the predictive behaviours of imported electricity flows are a first step in that direction (Maurice, 2015).
- Electricity consumption by the ICT network was not considered in this study byt should also be included in the scope of the study. Data transmission between data centres due to load balancing could be significant in the case of data storage services and result in non-negligible GHG emissions. If a GHG debt due to data traffic is attributed to each load migration, then each load migration should at least pay its debt. Such constraints could curb the number of migrations and thus limit the reduction in GHG emissions through load balancing. Also, the right emission factors to model emissions related to the network should be computed for each region that the ICT network crosses (assuming it is possible to map the network path taken by the data and attribute an electricity consumption property to each infrastructure of the ICT network). Research by Baliga et al. (2009) suggests that most of the electricity consumed by online data transmission via the Internet is attributed to the equipment used to access the network. Thus, a good approximation for data centres that rely on the Internet to transmit data would be to consider only the electrical mix of the region in which the data centre is located instead of considering all of the regions crossed by the ICT network.
- In this study, the focus was made to minimize marginal GHG emissions, but as mentioned previously, non-GHG emissions should also be considered to prevent impact displacements between impact categories.
- Data centres were considered to be similar in this study. However, if the cooling system of a data centre is considered in the calculations, differences in meteorological conditions could result in different energy consumptions to cool the servers. Therefore, weather forecasts and electricity needed to cool equipment might also be taken into account to manage loads between distributed data centres.
- In this study, the choice of a data centre that processes the load of the cloud computing service was the only parameter considered to minimize emissions. However, in the context of a Canadian smart grid and to have a fair assessment of the ICT footprint, many other parameters should be included in the optimization process. In fact, all electricity consumers (and producers) should play a role in minimizing global emissions. Especially, in the context of transport electrification (Velazquez et al., 2015), considering that electricity is a limited resource, it becomes pertinent to compare at a given time the environmental benefits of powering a data center (where the load is migrated to) or recharge batteries of electric vehicles even if these are two different domains. Thus, load migration between data centres is not the only measure to mitigate GHG emissions. Ideally, the management of marginal electricity should be based on all possible parameters, not only the ICT parameters.

4.6. Global applicability of the method

While the method followed in this paper is valid for many contexts, the results are specific to Ontario and Alberta for the period 2011-2013. These results should not be extrapolated to other regions without a proper adaptation. Indeed, the marginal electric mixes found for Ontario and Alberta may differ from those of regions presenting similar average electric mixes (e.g. France and Ontario due to a large contribution of nuclear power, or China and Alberta due to the large contribution of coal power).

Nevertheless, the main conclusions should remain valid in all contexts: marginal GHG emissions should be considered in real-time optimization systems, and GHG emissions accounting methods based on average emissions factors should be adapted to take into account marginal emission minimization. Moreover, because of the difficulty in identifying marginal electricity mixes, it would be appropriate for electric network operators to make these mixes public or to cooperate directly with cloud managers.

Concretely, the method can be used in every situation where it is intended to minimize GHG emissions in real-time by managing the power demand. Beyond optimization of ICT systems, this applies particularly to the management of smart-buildings electricity consumption, electric-vehicles battery load and any other type of demand side management programs and smart-grid projects.

4.7. Added value of the method

The conventional methods used to minimize in real-time the GHG emissions of data centre networks use average emissions factors of regional electricity generation. Application of the conventional methods result therefore in the exclusion of the emissions caused by the server load migrations. Consequently, the conventional methods might fail to optimize in real-time the management of server load migrations in order to minimize GHG emissions of data centre networks. In the worst situations, it is even possible the conventional methods drastically increase GHG emissions if fossil fuels are used to meet the power demand caused by server load migrations. By considering the marginal electricity, the proposed method improves the management of server load migrations used to minimize GHG emissions. Indeed, it enables a deeper vision of phenomena leading to GHG emissions. More generally, the proposed method improves decision taking by taking into account consequences of changes as compared to methods based on average emissions factors applied in the context of GHG emissions minimization of dynamic systems.

Conclusion

This paper highlights the need to take into account marginal electricity in the real-time optimization of multiregional data centre networks. It identifies an important issue regarding the conventional approach based on load balancing to minimize GHG emissions: the marginal electricity supplying the data centres that process the load after load migration is ignored. Therefore, the GHG emission reductions claimed by the conventional approach (if any) are uncertain. This study shows that the real-time management of load migrations between data centres located in different regions can lead to significant reductions in GHG emissions as compared to a single data centre solution. Interestingly, the identification of the marginal sources of electricity leads to GHG emission reductions that would not have been possible using the conventional approach based on average emission factors. Indeed, regional GHG emissions per kWh are quite different when considering marginal or average electricity. Results have been obtained for a Canadian case study involving two data centres in Ontario and Alberta. Each electric grid being unique, the computed emissions in this paper should not be directly extrapolated to other regions. However, the recommendation to consider marginal electricity in real-time optimization of multiregional data centre networks is expected to remain valid for regions using simultaneously several sources of electricity.

Beyond cloud computing and distributed data centres, the consideration for marginal emission factors appears also pertinent for smart systems (such as smart-building and smart-grids) where users adapt their power consumption to minimize their emissions.

However, minimizing marginal electricity emissions requires predictions of marginal sources of electricity, which are uncertain if the electric utilities do not make the information public in real time. Historic data may be used to built statistic models and predict marginal electricity mixes in real-time, but more research is needed to improve these models.

Moreover, the introduction of marginal electricity into the accounting of GHG emissions raises complex questions on the consideration of avoided marginal emissions. Such questions must be addressed in order to design a fair and efficient mechanism for the overall minimization and accounting of GHG emissions.

The optimization of distributed data centres is possible on the basis of GHG emissions reduction. However, it is not clear how the emissions of other undesirable substances are affected when GHG emissions are minimized. Studies must address this issue.

The optimization of the use phase of a cloud computing system should also consider the environmental impacts generated in the other life-cycle phases so as not to shift the impacts of the use phase to those phases. The manufacturing and end of life of ICT components are expected to have significant environmental impacts. Consequently, the environmental benefits through load migrations during the use phase within distributed data centres should at least compensate for the environmental impacts generated during the other life-cycle phases of these data centres.

Acknowledgments

The authors would like to thank NSERC and industrial partners Ericsson-Montreal and Inocybe Technologies for their financial support under grant CRDPJ 424371-11.

References

AESO. (2013). Market participant information - Data Requests. from http://www.aeso.ca/downloads/Hourly_Generation_by_Fuel.pdf

Amokrane, A., Zhani, M. F., Langar, R., Boutaba, R., & Pujolle, G. (2013). Greenhead: Virtual Data Center Embedding across Distributed Infrastructures. *Cloud Computing, IEEE Transactions on, 1*(1), 36-49. doi: 10.1109/TCC.2013.5

Amor, M. B., Pineau, P.-O., Gaudreault, C., & Samson, R. (2011). Electricity trade and GHG emissions: Assessment of Quebec's hydropower in the Northeastern American market (2006–2008). *Energy Policy, 39*(3), 1711-1721. doi: http://dx.doi.org/10.1016/j.enpol.2011.01.001

Arushanyan, Y., Ekener-Petersen, E., & Finnveden, G. (2014). Lessons learned – Review of LCAs for ICT products and services. *Computers in Industry, 65*(2), 211-234. doi: http://dx.doi.org/10.1016/j.compind.2013.10.003

Baliga, J., Hinton, K., Ayre, R., & Tucker, R. S. (2009). Carbon footprint of the internet. *Telecommunications Journal of Australia 59*(1), 1-5. doi: 10.2104/tja09005

Beloglazov, A., Abawajy, J., & Buyya, R. (2012). Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing. *Future Generation Computer Systems, 28*(5), 755-768. doi: http://dx.doi.org/10.1016/j.future.2011.04.017

Bonvoisin, J., Lelah, A., Mathieux, F., & Brissaud, D. (2012). An environmental assessment method for wireless sensor networks. *Journal of Cleaner Production, 33*(0), 145-154. doi: http://dx.doi.org/10.1016/j.jclepro.2012.04.016

Bonvoisin, J., Lelah, A., Mathieux, F., & Brissaud, D. (2014). An integrated method for environmental assessment and ecodesign of ICT-based optimization services. *Journal of Cleaner Production, 68*, 144-154. doi: http://dx.doi.org/10.1016/j.jclepro.2014.01.003

Borggren, C., Moberg, Å., Räsänen, M., & Finnveden, G. (2013). Business meetings at a distance – decreasing greenhouse gas emissions and cumulative energy demand? *Journal of Cleaner Production, 41*(0), 126-139. doi: http://dx.doi.org/10.1016/j.jclepro.2012.09.003

Buceti, G. (2014). Sustainable power density in electricity generation. *Management of Environmental Quality: An International Journal, 25*(1), 5-18. doi: 10.1108/meq-05-2013-0047

Dandres, T., Farrahi Moghaddam, R., Nguyen, K. K., Lemieux, Y., Cheriet, M., & Samson, R. (2014). Minimization of Telco cloud emissions based on marginal electricity management *ICT4S*. Stockholm, Sweden.

Deng, L., Babbitt, C. W., & Williams, E. D. (2011). Economic-balance hybrid LCA extended with uncertainty analysis: case study of a laptop computer. *Journal of Cleaner Production, 19*(11), 1198-1206. doi: 10.1016/j.jclepro.2011.03.004

Doyle, J., Shorten, R., & O'Mahony, D. (2013). Stratus: Load Balancing the Cloud for Carbon Emissions Control. *IEEE Transactions on Cloud Computing, 1*(1), 116-128. doi: 10.1109/TCC.2013.4

Farhat, A. A. M., & Ugursal, V. I. (2010). Greenhouse gas emission intensity factors for marginal electricity generation in Canada. *International Journal of Energy Research, 34*(15), 1309-1327. doi: 10.1002/er.1676

Farrahi Moghaddam, F. (2014). *Carbon-profit-aware job scheduling and load balancing in geographically distributed cloud for HPC and web applications.* École de technologie supérieure (ETS), University of Quebec (UduQ).

Farrahi Moghaddam, R., Farrahi Moghaddam, F., Dandres, T., Lemieux, Y., Samson, R., & Cheriet, M. (2014). *Challenges and complexities in application of LCA approaches in the case of ICT for a sustainable future*. Paper presented at the ICT4S, Stockholm, Sweden. http://atlantis-press.com/php/download_paper.php?id=13437

Frischknecht, R., & Rebitzer, G. (2005). The ecoinvent database system: a comprehensive web-based LCA database. *Journal of Cleaner Production, 13*(13-14), 1337-1343. doi: 10.1016/j.jclepro.2005.05.002

Gantz, J., & Reinsel, D. (2012). The Digital Universe in 2020 (pp. 1-16): IDC Analyse the Future.

Hawkes, A. D. (2010). Estimating marginal CO2 emissions rates for national electricity systems. *Energy Policy, 38*(10), 5977-5987. doi: http://dx.doi.org/10.1016/j.enpol.2010.05.053

IESO. (2013). Generator Output and Capability. 2013, from http://reports.ieso.ca/public/GenOutputCapability/

Jolliet, O., Margni, M., Charles, R., Humbert, S., Payet, J., Rebitzer, G., & Rosenbaum, R. K. (2003). IMPACT 2002+: A New Life Cycle Impact Assessment Methodology. *International Journal of Life Cycle Assessment, 8*(6), 7.

Kopsakangas-Savolainen, M., Mattinen, M. K., Manninen, K., & Nissinen, A. (2015). Hourly-based greenhouse gas emissions of electricity – cases demonstrating possibilities for households and companies to decrease their emissions. *Journal of Cleaner Production*. doi: http://dx.doi.org/10.1016/j.jclepro.2015.11.027

Krioukov, A., Goebel, C., Alspaugh, S., Chen, Y., Culler, D. E., & Katz, R. H. (2011). Integrating Renewable Energy Using Data Analytics Systems: Challenges and Opportunities. *IEEE Data Eng. Bull., 34*(1), 3-11.

Lannoo, B., Lambert, S., Van Heddeghem, W., Pickavet, M., Kuipers, F., Koutitas, G., . . . Aracil, J. (2013). Overview of ICT energy consumption (pp. 1-59): Network of Excellence in Internet Science.

Li, L., Hao, T., & Chi, T. (2016). Evaluation on China's forestry resources efficiency based on big data. *Journal of Cleaner Production*. doi: http://dx.doi.org/10.1016/j.jclepro.2016.02.078

Mandal, U., Habib, M., Shuqiang, Z., Mukherjee, B., & Tornatore, M. (2013). Greening the cloud using renewable-energy-aware service migration. *Network, IEEE, 27*(6), 36-43. doi: 10.1109/MNET.2013.6678925

Mathiesen, B. V., Münster, M., & Fruergaard, T. (2009). Uncertainties related to the identification of the marginal energy technology in consequential life cycle assessments. *Journal of Cleaner Production, 17*(15), 1331-1338. doi: DOI: 10.1016/j.jclepro.2009.04.009

Maurice, E. (2015). *Modélisation temporelle de la consommation électrique en analyse du cycle de vie, appliquée au contexte des TIC.* (Master), Université de Montréal, Montreal.

McCarthy, R., & Yang, C. (2010). Determining marginal electricity for near-term plug-in and fuel cell vehicle demands in California: Impacts on vehicle greenhouse gas emissions. *Journal of Power Sources, 195*(7), 2099-2109. doi: http://dx.doi.org/10.1016/j.jpowsour.2009.10.024

Meyer, D. E., & Katz, J. P. Analyzing the environmental impacts of laptop enclosures using screening-level life cycle assessment to support sustainable consumer electronics. *Journal of Cleaner Production*. doi: http://dx.doi.org/10.1016/j.jclepro.2015.05.143

Meza, J., Shih, R., Shah, A., Ranganathan, P., Chang, J., & Bash, C. (2010, November 12-18, 2010). *Lifecycle-based Data Center Design.* Paper presented at the ASME 2010 International Mechanical Engineering Congress & Exposition, Vancouver, British Columbia, Canada.

Mirabella, N., Rigamonti, L., & Scalbi, S. (2013). Life cycle assessment of Information and Communication Technology application: a case study of dematerialization in the Italian Public Administration. *Journal of Cleaner Production, 44*(0), 115-122. doi: http://dx.doi.org/10.1016/j.jclepro.2012.10.051

Nielsen, S., Sorknæs, P., & Østergaard, P. A. (2011). Electricity market auction settings in a future Danish electricity system with a high penetration of renewable energy sources – A comparison of marginal pricing and pay-as-bid. *Energy, 36*(7), 4434-4444. doi: http://dx.doi.org/10.1016/j.energy.2011.03.079

Olkkonen, V., & Syri, S. (2016). Spatial and temporal variations of marginal electricity generation: the case of the Finnish, Nordic, and European energy systems up to 2030. *Journal of Cleaner Production, 126*, 515-525. doi: http://dx.doi.org/10.1016/j.jclepro.2016.03.112

Raichur, V., Callaway, D. S., & Skerlos, S. J. (2016). Estimating Emissions from Electricity Generation Using Electricity Dispatch Models: The Importance of System Operating Constraints. *Journal of Industrial Ecology, 20*(1), 42-53. doi: 10.1111/jiec.12276

Rogers, M. M., Wang, Y., Wang, C., McElmurry, S. P., & Miller, C. J. (2013a). Evaluation of a rapid LMP-based approach for calculating marginal unit emissions. *Applied Energy, 111*, 812-820. doi: http://dx.doi.org/10.1016/j.apenergy.2013.05.057

Rogers, M. M., Wang, Y., Wang, C., McElmurry, S. P., & Miller, C. J. (2013b). Evaluation of a rapid LMP-based approach for calculating marginal unit emissions. *Applied Energy, 111*(0), 812-820. doi: http://dx.doi.org/10.1016/j.apenergy.2013.05.057

Roux, C., Schalbart, P., & Peuportier, B. (2016). Accounting for temporal variation of electricity production and consumption in the LCA of an energy-efficient house. *Journal of Cleaner Production, 113*, 532-540. doi: http://dx.doi.org/10.1016/j.jclepro.2015.11.052

Şengül, H., Bayrak, F., Aydınalp Köksal, M., & Ünver, B. (2016). A cradle to gate life cycle assessment of Turkish lignite used for electricity generation with site-specific data. *Journal of Cleaner Production, 129*, 478-490. doi: http://dx.doi.org/10.1016/j.jclepro.2016.04.025

Sensfuß, F., Ragwitz, M., & Genoese, M. (2008). The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. *Energy Policy, 36*(8), 3086-3094. doi: http://dx.doi.org/10.1016/j.enpol.2008.03.035

Siler-Evans, K., Azevedo, I. L., & Morgan, M. G. (2012). Marginal Emissions Factors for the U.S. Electricity System. *Environmental Science & Technology, 46*(9), 4742-4748. doi: 10.1021/es300145v

Société en commandite Gaz Métro, Association des consommateurs industriels de gaz, Corporation des propriétaires immobiliers du Québec, Fédération canadienne de l'entreprise indépendante, Groupe de recherche appliquée en macroécologie, Option consommateurs, . . . Union des municipalités du Québec. (2007). Mécanisme incitatif *Renouvellement du mécanisme incitatif à l'amélioration à la performance de Gaz Métro (R-3599-2006)* (pp. 1-60). Montreal: Régie de l'énergie du Québec.

The Climate Group. (2008). SMART 2020: Enabling the low carbon economy in the information age (pp. 1-87): Global eSustainability Initiative.

Thomas, C. E. (2012). US marginal electricity grid mixes and EV greenhouse gas emissions. *International Journal of Hydrogen Energy, 37*(24), 19231-19240. doi: http://dx.doi.org/10.1016/j.ijhydene.2012.09.146

Tveten, Å. G., Bolkesjø, T. F., Martinsen, T., & Hvarnes, H. (2013). Solar feed-in tariffs and the merit order effect: A study of the German electricity market. *Energy Policy, 61*(0), 761-770. doi: 10.1016/j.enpol.2013.05.060

Van Heddeghem, W., Vereecken, W., Colle, D., Pickavet, M., & Demeester, P. (2012). Distributed computing for carbon footprint reduction by exploiting low-footprint energy availability. *Future Generation Computer Systems, 28*(2), 405-414. doi: 10.1016/j.future.2011.05.004

Velazquez, L., Munguia, N. E., Will, M., Zavala, A. G., Verdugo, S. P., Delakowitz, B., & Giannetti, B. (2015). Sustainable transportation strategies for decoupling road vehicle transport and carbon dioxide emissions. *Management of Environmental Quality: An International Journal, 26*(3), 373-388. doi: 10.1108/meq-07-2014-0120

Whitehead, B., Tozer, R., Shah, A., Kosik, B., Maidment, G., Andrews, D., & Dunn, A. (2012). *The Environmental Burden of Data Centres – A Screening LCA Methodology.* Paper presented at the CIBSE ASHRAE Technical Symposium, Imperial College, London UK.

Yang, C. (2013). A framework for allocating greenhouse gas emissions from electricity generation to plug-in electric vehicle charging. *Energy Policy, 60*, 722-732. doi: http://dx.doi.org/10.1016/j.enpol.2013.05.013

Zheng, Z., Han, F., Li, F., & Zhu, J. (2015). Assessment of marginal emissions factor in power systems under ramp-rate constraints. *CSEE Journal of Power and Energy Systems, 1*(4), 37-49. doi: 10.17775/CSEEJPES.2015.00049