



**HAL**  
open science

# A generic GIS-based method for small Pumped Hydro Energy Storage (PHES) potential evaluation at large scale

Antoine Rogeau, Robin Girard, Georges Kariniotakis

► **To cite this version:**

Antoine Rogeau, Robin Girard, Georges Kariniotakis. A generic GIS-based method for small Pumped Hydro Energy Storage (PHES) potential evaluation at large scale. *Applied Energy*, 2017, 197, pp.241 - 253. 10.1016/j.apenergy.2017.03.103 . hal-01513139

**HAL Id: hal-01513139**

**<https://minesparis-psl.hal.science/hal-01513139>**

Submitted on 24 Apr 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A generic GIS-based method for small Pumped Hydro Energy Storage (PHES) potential evaluation at large scale

A. Rogeau<sup>a,\*\*</sup>, R. Girard<sup>a,\*</sup>, G. Kariniotakis<sup>a</sup>

<sup>a</sup>MINES ParisTech, PSL Research University, PERSEE - Centre for processes, renewable energy and energy systems, CS 10207 rue Claude Daunesse 06904 Sophia Antipolis Cedex, France

---

## Abstract

The increasing share of weather-dependent renewable energies in power systems creates a need for energy storage technologies to reduce the impacts of variable production. The most mature technology to store energy on the grid remains Pumped Hydro Energy Storage (PHES). The potential of high-energy sites has already been assessed in Europe by the EU JRC, considering mostly dams and reservoirs from global European databases which include only massive water bodies. This paper focuses on estimating the potential for small-PHES, proven to have lower environmental impact and a positive impact on grid balance and reliability. A generic method is designed, able to evaluate a global PHES storage capacity at large scale. It considers both existing lakes and natural depressions suitable to be filled for PHES purposes. The volume of filled lakes is estimated using the surrounding topography. The method is organized so that the "heavy" calculations, i.e. sink detection, volume evaluation, constraints verification etc. are run only once. Consequently, the actual potential estimation phase only includes fast calculations and can be integrated in a loop for carrying out a sensitivity analysis. The proposed method is then applied considering France as a test case. Suitable environmental, land-use and structural constraints are applied to eliminate irrelevant sites. The analysis leads to an estimated value of the small-PHES potential in France, which ranges from 14 GWh when only existing lakes are considered to 33 GWh when lakes and depressions are considered. These estimations represent respectively 8% and 18% of the current hydro storage capacity in France. Thanks to a global sensitivity analysis, factors like the maximum distance between lakes, the maximum altitude of the sites, and the distance to the electrical grid are shown to have the most influence on the global evaluated potential, which is further sensitized. Lastly, another application is suggested that makes it possible to select the connections to be built first within a restricted area, based on a cost-per-energy-like approach. It uses the connections between reservoirs detected at large geographical scale.

**Keywords:** Pumped hydro energy storage, GIS, Potential, Sobol analysis, Renewable energies, France

---

## 1. Introduction

Over the last few decades, renewable energy sources (RES) have continuously increased their share in the world energy market. In fact, worldwide RES installed capacity went from 800 GW in 2004 - mostly from hydropower sources (715 GW) - to almost 1850 GW by the end of 2015 [1] - 1064 GW from hydropower. During this period, the capacities of wind power (WP) and solar photovoltaic (PV) installations have multiplied by thirteen, going from 50 to 660 GW [1]. Both PV and WP production are highly variable and not entirely dispatchable due to their inherent dependence on weather conditions. To handle RES production fluctuations, different solutions exist, such as demand management, the combination of RES plants with other means of conventional generation, and also energy storage. Several storage technologies exist [2], including electrochemical devices such as batteries or fuel cells and mechanical solutions such as pumped hydro energy storage (PHES).

Water lifting devices have been invented, used and improved by humans for thousands of years [3]. Following the evolution of these technologies, PHES is now one of the most mature and cost-effective solutions for absorbing RES power fluctuations. Rehman et al. [4] and Uniyal [5] estimated the global pumped storage installed capacity as slightly more than 100 GW in 2010.

Large PHES facilities are usually employed as grid-scale solutions to ensure that electricity production matches demand at any time, especially thanks to their fast response. PHES offers considerable storage capacities at affordable prices [6]. In Europe, high-potential sites for large PHES have been assessed by an EU JRC study [7], which considered artificial reservoirs larger than 100,000 m<sup>3</sup> separated by at least 150 m of head. This resulted in a total European feasible potential varying from 36 to 4000 GWh for reservoirs at a distance of between 3 and 20 km. For France, this potential reaches 4 GWh if the distance between lakes is limited to 3 km and rises to 500 GWh at 20 km. Large PHES schemes were initially considered as the only viable type of pumped storage due to elevated inversion costs and scale effects [6]. However, small pumped hydro energy storage (small-PHES) has recently been studied more deeply. Small-PHES has a lower environmental

---

\* Principal corresponding author

\*\* Corresponding author

Email addresses: antoine.rogeau@mines-paristech.fr (A. Rogeau), robin.girard@mines-paristech.fr (R. Girard), georges.kariniotakis@mines-paristech.fr (G. Kariniotakis)

impact in terms of CO<sub>2</sub> equivalent [8]. Large-PHES has a national and continental role to play, while the role of small-PHES is local and regional. The use of local small-PHES utilization can reduce peak demand and thus avoid or defer electrical grid reinforcement [9, 10]. Small-PHES can also mitigate voltage issues in medium and low voltage lines, due to the injection of decentralized production [11]. In brief, small-PHES can contribute to the decentralized integration of intermittent RES. In this perspective, PHES systems have been miniaturized and are even starting to be used in cities, installed on top of buildings [12]. The present paper thus focuses on evaluating the potential of sites connecting close reservoirs of limited volume that are suitable for installing small-PHES plants. This method does not aim to reduce the PHES costs in itself. It aims to find spots requiring few civil work and thus exploitable with little expenditure. For this purpose, a new computationally efficient methodology applicable at large scale (i.e. national) is proposed. The method employs Geographic Information Systems (GIS) to detect reservoirs, associate those that could host a small-PHES plant, and finally apply the different constraints to derive a feasible potential. One disadvantage of small pumped hydro energy storage is the investment cost, given the low storage capacity. Small PHES indeed suffers from scale effects [13]. The method developed considers natural opportunities, such as existing lakes and natural terrain depressions, which could be operated at a limited cost. The proposed methodology is applied to the case of France to evaluate small-PHES potential at national scale.

For both the conventional and pumped hydropower systems, the location of the power plant is important, as the extracted energy and power depend, among other factors, on the gross head, water flow and volume of water available. Selecting optimal spots that respect these parameters thus appears to be crucial. For this purpose, automatic numeric methods have been developed to select the best spots, taking advantage of the emergence of GIS. Thereby, Larentis et al. [14] present a method to detect the optimal location in a river basin to install a run-of-river hydropower plant. They use GIS to detect where to install a dam and build the powerhouse. Kusre et al. [15] evaluate suitable places in an Indian valley, using hydrological tools and soil, land use and weather data. Both focus on power generation, without storage and optimize the location within restricted areas, mostly valleys.

Other studies concern the PHES scheme, and are carried out in limited areas to evaluate either the potential of a zone, or a particular spot to install a PHES plant. Ahmadi and Shamsai [16] focus on a single massive lake, looking for the best location to construct the corresponding reservoir using criteria ranking through raster analysis. This is a localized analysis, with an emphasis on very detailed geological or environmental data, which makes application of this method at a larger scale difficult. Mailler et al. [17] focus on small hydropower schemes, taking advantage of existing reservoirs in Switzerland, such as artificial snow reserves. GIS data are not used and the study is conducted in a very restricted area. Kucukali [18] estimates the best location for conversion of existing hydropower plant to PHES schemes comparing a few sites in

Turkey, with ranking based on geological, social and environmental criteria. Garcia [19] focuses on one very restricted region, looking for PHES opportunities within a group of many lakes. Several potential connections were identified and the selection was then optimized depending on criteria such as stored energy potential, and the cost of stored energy potential.

A common characteristic of these four studies is that they are conducted on a small scale and based on very detailed datasets (geological composition, ground slopes, urban planning, etc.), which are difficult to obtain for a large area.

Other works are based on the detection of high power (and energy) capacity PHES sites. Connolly et al. [20] proposes a method that exploits a Digital Elevation Model (DEM) to identify restructurable hills and build reservoirs in the south of Ireland. This method implies considerable investment costs, which are only profitable if they correspond to high-power projects. Four major works study energy storage possibilities at a larger scale [21, 22, 7, 23]. The first one is focused on taking advantage of existing dams, but also includes a section on matching existing lakes in the United States. It is oriented to large PHES schemes, but only focuses on specific sites and does not offer a global potential evaluation in terms of energy or power. Similarly, the second study only searches for high-power locations throughout Turkey, with high head or volume characteristics (150 m and more than 10<sup>6</sup> m<sup>3</sup>). It also includes a search for flat locations to dig new reservoirs. Likewise, the third document is a global study of potential PHES sites. However, it includes a European assessment of energy available per country, depending on a maximum authorized distance between reservoirs of 3 to 20 km. Only heads greater than 150 m are considered. Lu and Wang [23] present an evaluation of Tibet's potential for PHES. It considers connection with very high heads (greater than 500 m) and large reservoir areas (greater than 60,000 m<sup>2</sup>).

All the above works exclude connections with low head and low potential, which might be beneficial for small-pumped storage due to restricted civil works, as we aim to evaluate here.

Pauwels et al. [24] conducted a study at the scale of a French county (about 4000 km<sup>2</sup>). It matches natural terrain depressions within a wide constrained zone. Existing lakes are not taken into account. Moreover, no optimization through interconnected lakes is performed, thus leading to an overestimation of energy potential. The evaluation by Gimeno-Gutierrez and Lacal-Artegui [7] presents a substantial shortfall in the potential reservoir dataset, by considering only existing artificial reservoirs.

Considering only a share of potential reservoirs can result in a considerable shortfall in evaluations. Moreover, in order to avoid matching one lake with several others, it is necessary to select the most *effective* connections, which results in more precise estimations of the overall potential. In Fitzgerald et al. [22], a sensitivity analysis is carried out featuring the buffer size, minimum head, and maximum slope of flat terrains compatible with PHES installations. Nonetheless, each criterion is treated separately, which makes it difficult to understand the

global behaviour of the output. Similarly, Garcia [19] presents a sensitivity analysis of the evaluated potential, drawing an abacus representing the evolution of the detected energy depending on storage duration and distance between reservoirs. It clearly identifies the latter factor as more decisive. However, the analysis is only carried out for two input parameters. These sensitivity analyses remain focused on a limited number of parameters and are not globally applied to all of the parameters of the corresponding evaluation method.

The main contribution of this paper is to propose a generic method able to evaluate the potential of small-PHES over a large geographical zone (several hundreds or thousands of square kilometres). To ensure that a significant share of potential reservoirs are considered, the method includes the dataset of both existing water bodies and natural terrain depressions. It can be applied at a large scale thanks to the global datasets selected. Connections compatible with small-PHES purposes (e.g. not too massive) are evaluated, thus opening up a new perspective in comparison with former studies. A pairing of potential reservoirs is carried out authorizing a maximum of one connection per reservoir, to increase the accuracy of the evaluation. Moreover, the method includes an optimization step to select the most appropriate connections when a lake is concerned by various matchings. A global Sensitivity Analysis is incorporated into the method and is feasible thanks to the low execution time of the evaluation process. As a second contribution, the developed method is applied to the whole of France (about 675,000 km<sup>2</sup>). This paper presents the overall estimated potential but also the spatial distribution of detected connections. The proposed global SA is applied to the same case study. Additionally, a potential avenue for further development is to use the obtained connections database as an alternative way of making a best site selection analysis.

## 2. Method description

The proposed methodology consists in identifying existing reservoirs - either natural water bodies (referred to as lakes) or sinks in the topography (referred to as depressions) - and searching for another compatibly located potential reservoir nearby in order to reference them as potential small-PHES sites. The method is divided into four main parts.

1. Reservoir detection: (See 2.1)
  - a) Existing lakes: Estimation of volume.
  - b) Potential new reservoir sites: Relief treatment for the detection and characterization of natural terrain depressions.
2. Reservoir connections creation: (See 2.2)
  - a) Matching of suitable reservoir sites: Creation of a large dataset with potential pairings.
  - b) Verification of the acceptability constraints on paired reservoirs and identified connections with environmental preservation zones or remarkable land-use areas
3. Constraints application: (See 2.3)
  - a) Application of an acceptability filter: No-go zones defined for excluding pairings in sensitive areas

- b) Application of a structural filter: maximum distance between lakes, minimum head, maximum volume, etc. defined by the user for the potential evaluation
4. Optimization of selected connections: (See 2.4)
  - a) Optimization of the pairs resulting from step 3: one connection per lake, maximizing energy.
5. Sensitivity Analysis: (See 2.5)
  - a) Impact of sensitive zones considered or not as No-go (step 3) on the evaluated potential.
  - b) Application of a Sobolj sensitivity analysis method to detect factors impacting the output

This methodology is summarized by Figure 1.

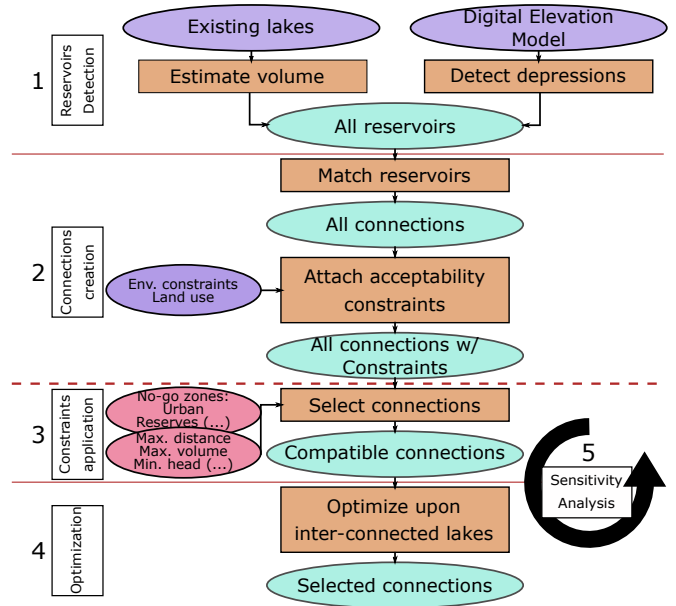


Figure 1: Methodological flowchart evaluating the small-hydro pumped storage capacity at large scale.

### Main model assumptions

The proposed approach considers only closed-loop systems, excluding rivers and the sea as a lower reservoir. Only connections between two inland water-bodies are considered.

#### 2.1. Reservoir detection

In a simple closed-loop hydro storage scheme, two water bodies are used to pump water up and down, in order to store and generate electricity. The primary type of reservoir suitable for this purpose is existing lakes. Given that they are already filled with water and hermetic, they constitute the easiest resource for installing a PHES plant.

##### 2.1.1. Existing lakes

Databases containing water body information on a large scale - such as the Global Reservoir and Dam (GRAND), the European catchments and Rivers network system (Ecrins) or the DAM Positioning (DAMPOS) databases - focus on very large reservoirs. For smaller water bodies, national institutes

are better suited, such as NGI in Belgium, SNIG in Portugal and IGN in France, all of which provide appropriate national data-bases.

These databases, which are often supported by GIS, generally include the water bodies, their altitude, regime and nature, but no information on volume. However, this information is essential for hydro storage because it defines the energy available in a connection. Some studies have been carried out to estimate reservoir volume using surface area as the only input. Liebe et al. [25] and Sawunyama et al. [26] developed methods estimating the volume of a reservoir as a function of its surface area in Ghana and Zimbabwe. Nonetheless, both works concerned small reservoirs and required manual categorization of lakes. In this work, we evaluate the volume of lakes using the method proposed in Hollister et al. [27], by which we can estimate a natural WB's volume from its surrounding topography. This approach is computationally efficient for the large-scale study we are aiming at, compared to alternatives that would give an exact estimation of the volume. Altitude information can be found using global Digital Elevation Models (DEM) and especially the SRTM model, available worldwide and free of charge. National institutes also often possess their own national DEM. Typically, the spatial resolution of such models is 25 m.

### 2.1.2. Natural depressions

Beyond existing lakes, the second objective of this step is to identify potential reservoirs by evaluating natural terrain depressions that might be flooded. The same DEM as in 2.1.1 can be used to carry out this step.

First, all depressions in the DEM must be detected. The methodology described by Planchon and Darboux [28] allows us to obtain all terrain depressions, of any size and on any territory. To ensure that very large valleys are not included in the research, we can set a maximum surface area. A threshold of 1 square kilometre seems reasonable for such small-PHES schemes. Depressions exceeding this limit are then scanned downwards to identify smaller internal depressions. Volume calculation is straightforward. The smaller sinks can be eliminated using a minimum energy criterion (20 kWh/m for example, corresponding to Volume <8000 m<sup>3</sup>).

## 2.2. Connection creation

### 2.2.1. Reservoir matching

One of the objectives of the proposed method is its computational efficiency. To achieve a fast simulation, the matching detection is achieved following two major aspects:

First, a database is created, containing all potential small-PHES connections. To do so, two settings are defined, regarding the minimum head and the maximum distance between lakes (See Table 1). These are chosen at not very restrictive values so that a maximum amount of pairs are included in the dataset. Concretely, a reservoir will be paired with every other reservoir presenting a head of minimum 10 m within a representative diameter of 3 500 m. For each pair of lakes that satisfies the above criteria, the following parameters are estimated using the available terrain databases and then included

Table 1: Basic settings for the connections dataset building.

Maximum distance between reservoirs	Minimum head between reservoirs
3 500 m	10 m

in the database with all of the potential small-PHES connections:

- Distance between reservoirs (m)
- Gross head between reservoirs (m)
- Energy of the connection (kWh)
- Volume of the smallest reservoir - available to transfer (m<sup>3</sup>)
- Depth of the shallowest reservoir (m)
- Altitude of the upper reservoir (m)
- Shortest distance to the electrical grid (m)

The energy of the connection is calculated as follows:

$$E = \eta\rho gHV \quad (1)$$

where  $\eta$  is the efficiency of the pumped storage plant (chosen as 0.7 in our case, the typical efficiency of small-PHES [29]),  $\rho$  is the water density (kg/m<sup>3</sup>) and  $g$  the gravity acceleration (m/s<sup>2</sup>).  $H$  is the altitude difference between the bottom of the upper reservoir and the surface of the lower reservoir, called the gross head of the connection (m).  $V$  is the available water volume to transfer (m<sup>3</sup>), calculated as 90% of the volume of the smaller of the two involved reservoir.

The shortest distance to the electrical grid is the distance from the lower reservoir to the high- or medium- voltage electrical grid. This electric network database, generally produced by the national geographic institute or the grid owner itself, has to be imported for the study zone. Installing a power plant very far from the existing network would make it difficult to carry out the project, as it adds costs and involves civil works.

### 2.2.2. Attaching acceptability constraints information

Some of the potential reservoirs detected, i.e. existing lakes and natural terrain depressions, might be ineligible for PHES application. Many zones are preserved from human presence and constructions to protect fauna and flora. Many kinds of preservation zones exist but they do not all ban energy-related constructions. The World Database Protection Area (WDPA)<sup>1</sup> provides an inventory of these protected zones worldwide. These zones can be regulated by regional, national or international instances, and involve different restrictions.

Additional constraints that can lead to the exclusion of some locations relate to land use. These may include urban

<sup>1</sup><http://blog.protectedplanet.net/>

areas, leisure areas (i.e. navigation, swimming, etc.) or professional activity zones (i.e. flooding arable land might make a project difficult to complete). Available land-use datasets differ depending on the study zones, but a good option for Europe is the Corine Land Cover<sup>2</sup>. The 2015 publication relating to 2012 land cover is available for the whole of Europe and describes land use, dividing activities into several sections. Whatever the dataset, the following five categories of land use can be considered as relevant:

- **Urban zones:** Most of the land is covered by a water-proofed surface. Most of it features buildings or roads.
- **Leisure activities:** The land is used for camping, sports and other leisure activities.
- **Arable land:** Land capable of being ploughed and used to grow crops.
- **Permanent farming:** Permanent farming land, including crops such as orchards, vineyards etc.
- **Pasture:** Mostly land used to feed livestock.

Moreover, one of the main reasons for eliminating a natural depression from potential PHES sites is when it features human constructions. The detection of any building within natural depressions must be verified using a building database (obtained from national institutes for example). Each of the constraints cited above is applied to the whole set of small-PHES compatible connections, in order to verify whether the connected reservoirs are present in a sensitive area. If so, the connection involving these reservoirs could be discarded in an upcoming process. Further evaluation of the storage energy available across the study zone is then carried out, depending on which of these acceptability constraints are considered to be incapacitating. Secondly, for each detected connection, the inclusion of the future pipe, represented by a line between the two matched reservoirs, is verified in line with the constraints concerning urban zones.

Consequently, the three following pieces of acceptability information are attached to the detected connection pairs:

- Inclusion of matched reservoirs in land-use zones, environmental protection areas (vector with binary values 0 or 1)
- Presence of buildings in any reservoir (0,1 vector)
- Inclusion of the pipe in a urban zone (0,1 vector)

Finally, the detected connection pairs are divided into three main groups, defined as shown in Table 2.

Table 2: Denomination of connections depending on the nature of connected reservoirs

Name	Initials	Description
Type 1	T1	Pairs composed of two existing reservoirs
Type 2	T2	Pairs composed of one existing lake and one terrain depression detected by the DEM method
Type 3	T3	Pairs composed of two terrain depressions detected by the DEM method

### 2.3. Constraints application

Among the numerous connection pairings detected, some have to be discarded using the parametrized constraint data generated in 2.2.2 and before the next step of optimization. Initially, a number of pairings is eliminated after being checked against the land-use constraints. As a default configuration, restrictions on such constraints are suggested in Table 3. Likely

Table 3: Ranking of the different constraints, environmental and land-use

Restriction	Ranking
Urban zones	No-go
Leisure activities	No-go
Arable land	No-go
Permanent farming	No-go
Pasture	Suitable

environmental preservation rankings must be defined depending on the preservation area dataset used for the study zone. Generally, international preservation zones and natural reserves can be considered as No-go zones.

Connection pairings with configurations that do not correspond to a small-PHES scheme also have to be eliminated. For this purpose, nine main criteria are selected; their respective thresholds are defined in Table 4. We choose a minimum

Table 4: Limitation for different connection parameters

Criterion	Limit	Default value
Head	Minimum	40 m
Distance	Maximum	2 500 m
Volume	Minimum	0 m <sup>3</sup>
Volume	Maximum	750,000 m <sup>3</sup>
Depth	Minimum	0 m
Grid distance	Maximum	10,000 m
Altitude	Maximum	2 500 m
Energy	Minimum	5000 kWh
Head/Distance ratio	Minimum	0.1

energy threshold so that the minimum sizing accepted corresponds to a 500 kW plant active for 10 hours. The minimum Head/Distance ratio corresponds to a 10% slope [30] [23].

<sup>2</sup><http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012/view>

#### 2.4. Optimization of inter-connected pairings

Among the selected connection pairings, some may include reservoirs featuring in more than one pairing. We consider that the small-PHES scheme can involve only two lakes (one reservoir can not be linked to several others). An optimization is done as explained below to obtain non-interconnected pairings.

The problem to solve can be seen as a multidimensional knapsack problem [31], defined as follows:

$$\max Z = \sum_{j=1}^n c_j x_j \quad (2)$$

$$\text{subject to } \sum_{j=1}^n a_{ij} x_j \leq 1, \quad \forall i \in \llbracket 1, m \rrbracket \quad (3)$$

$$x_j \in \{0, 1\}, \quad \forall j \in \llbracket 1, n \rrbracket \quad (4)$$

where  $j$  is the connection number and  $i$  is the lake number.  $x_j$  is the binary representing whether connection  $j$  must be built or not and  $c_j$  is the "benefit" of the connection. Equation (3) represents the fact that one lake  $i$  can be paired only once, where  $a_{ij}$  is a matrix describing whether connection  $j$  concerns lake  $i$ . Equation (4) describes  $x_j$  as a binary.

As a result, connections are selected so that each reservoir is connected no more than once. In some cases, more than one connection can remain within a group of interconnected reservoirs (if there are more than 3 reservoirs).

#### Benefit definition

Different benefit vectors can be defined before running the optimization tool:

- Energy, as calculated in (1). Maximizing the energy would provide the biggest storage capacity in the study zone.
- Power, calculated as  $P = \rho g H Q \eta$ , where  $Q$  is the flow rate in the pipe ( $\text{m}^3/\text{s}$ ). This strategy implies defining the maximum water flow for each connection and favours the pairs with high head, which provide bigger power.
- Cost-energy ratio, calculated as:

$$CoE = \frac{I}{E} \quad (5)$$

where  $I$  is the investment costs and  $E$  the energy available in the connection.

This cost, expressed in €/KWh, allows us to choose more expensive connections if they offer a considerable energy advantage. This strategy also creates the possibility of eliminating the least effective connections (below a defined threshold).

- Other functions, such as levelized cost of energy, calculated as follows:

$$LCOE = \frac{\sum_{t=1}^n \frac{I_t + M_t}{(1+r)^t}}{\sum_{t=1}^n \frac{E_t}{(1+r)^t}} \quad (6)$$

where  $I_t$  is the investments costs at year  $t$ ,  $M_t$  is the operation and maintenance costs at year  $t$ ,  $E_t$  is the benefit (here energy or power) in the year  $t$ , and  $r$  is the discount rate.

- Other benefit functions can be used, such as cost of power.

In case of using energy or power as a benefit vector, the above optimization problem can be adapted to give priority to solutions including existing lakes, which are generally less expensive to fill than depressions. This translates into selecting T1 connections first, then T2, and ultimately the remaining T3 connections.

#### 2.5. Sensitivity analysis

##### 2.5.1. Impact of the defined exclusion zones

In section 2.3, connection pairings are eliminated from the evaluation because of their inclusion in constraint zones defined by the user as No-go. The binary state "No-go/Suitable" is generally not so clear in reality, and the impact of this uncertainty can be significant on a large study area. For this reason, it is recommended to identify the sensitive constraint areas that are most likely to be between the No-go and suitable status, and evaluate the variation of the energy storage capacity whether these specific constraints are accepted or not. A quantification of the output variation around a default state should be achieved, constraint by constraint.

##### 2.5.2. Sensitivity analysis method

As described in Padey et al. [32], a sensitivity analysis (SA) is an approach for investigating results variability as a function of the input parameters. Most SAs only consider variation around a nominal value of one parameter at a time, or from a worst case to best case scenario. This results in mis-estimations of the influence of each parameter on the function output. The Global Sensitivity Analysis (GSA) proposed by Sobol [33] overcomes these limitations. Moreover, this model does not require linear modelling assumptions.

The Sobol indices reflect the shares of variance in the output that are explained by variations in the input parameters. The higher the value of a Sobol index, the higher the impact of the corresponding parameter on the model output. The overall methodology is derived from Sobol [33] and is divided into 4 sub-steps:

- Stochastic generation of  $n$  scenarios,
- Computation of the impacts, applying the method with the  $n$  scenarios as inputs,
- Estimation of the Sobol indices for each parameter,
- Selection of the key parameters explaining most of the variability.

### 2.5.3. Sobol indices definition

To quantify the sensitivity of the resource estimation with respect to the various inputs considered in our model, we choose to use Sobol indices [33]. These are global sensitivity indices designed to estimate the relative share of output variance determined by the considered inputs or their interactions. Sobol indices are very popular because their computation does not rely on any (e.g. linear) hypothesis about the underlying model function  $f$  that links input  $X_i$ ,  $i \in \phi := \{1, \dots, p\}$  to output  $Y = f(X_\phi)$ . However, they assume that inputs are uncorrelated. Then, the following generic decomposition also has uncorrelated terms:

$$Y = \mathbb{E}(Y) + \sum_{I \subseteq \phi} f_I(X_I)$$

$$\text{with } f_I(X_I) = \mathbb{E}[Y|X_I] - \sum_{I' \subseteq I} f_{I'}(X_{I'}) \quad (7)$$

which allows us to decompose the variance  $V$  of  $Y$  in the same way. Normalizing by  $V$  finally generates Sobol indices  $S_I$ :

$$1 = \sum_{I \subseteq \phi} S_I, \text{ with } S_I = V_I/V \text{ and } V_I = \text{var}(f_I(X_I)) \quad (8)$$

Various summary statistics based on Sobol indices may then be proposed. In Homma and Saltelli [34], the total sensitivity of a system  $S_{T_i}$  to a given parameter  $X_i$  is defined by adding up all indices that involve parameter  $X_i$ , i.e.  $S_{T_i} = \sum_{I, i \in I} S_I$ . Extending that idea, it may be interesting to evaluate the total sensitivity of a system  $S_{T_L}$  with some general aspect (e.g. technological characteristics, relief constraints, etc.) whose definition potentially involves several parameters  $L \subseteq \phi$ , i.e.

$$S_{T_L} = \sum_{I, I \cap L \neq \emptyset} S_I \quad (9)$$

### 2.5.4. Sobol indices estimation

First, note that thanks to the definition of  $f_I(X_I)$  in Equation 7 we can recursively estimate Sobol indices, i.e.

$$S_I = \frac{\text{var}(\mathbb{E}[Y|X_I])}{V} - \sum_{I' \subsetneq I} S_{I'} \quad (10)$$

Thus, to estimate  $S_I$ , given that we are only considering first order indices, we only need to estimate  $\text{var}(\mathbb{E}[Y|X_I])$  (and  $V$ ). This can be done through Monte Carlo simulations using two independent, equally sized, random samples of inputs:  $(X_\phi^{j,1})$  and  $(X_\phi^{j,2})$ , where  $j = 1, \dots, n$  [34]:

$$\hat{f}_0 = \hat{\mathbb{E}}(Y) = \frac{1}{n} \sum_{j=1}^n f(X_\phi^{j,(1,2)}) \quad (11)$$

$$\hat{V} = \frac{1}{n} \sum_{j=1}^n f^2(X_\phi^{j,(1,2)}) - \hat{f}_0^2 \quad (12)$$

$$\hat{\text{var}}(\mathbb{E}[Y|X_I]) = \frac{1}{n} \sum_{j=1}^n f(X_\phi^{j,1})f(X_\phi^{j,2}) - \hat{f}_0^2 \quad (13)$$

where  $\bar{I} = \phi \setminus I$ . Estimating Sobol indices requires  $(K+1)n$  evaluations of  $f$ ,  $K$  being the number of considered indices.

## 3. Application of the method to France

The method described in Section 2 is applied to a whole country, i.e. France (Area: 675,000 km<sup>2</sup>). This section presents the software employed to implement the method, the datasets used for this purpose and the results of the evaluation. A very local study case is presented to demonstrate the efficiency of the optimization tool.

### 3.1. Dataset choice and connection detection

Existing water body data is provided by the French National Geographic Institute (IGN) TOPO database. This contains all WBs over 20 metres long detected by satellite. It represents about 820,000 lakes, rivers and ponds. Rivers are excluded.

The DEM data, used to evaluate the volume of lakes and detect sinks, comes from the IGN ALTI database, offering 25 m x 25 m resolution of the whole of France. Due to the large size of the territory to process, the whole dataset (WB and DEM) was re-organized into a 15 m x 15 m-cell grid. This represents 138 land rasters measuring about 70 km down one side. To be certain that all depressions and further connections are taken into account, even between raster-cells, the original rasters were extended with 5-kilometre buffers, on the right and bottom sides, as shown in Figure 2. To confirm the ef-

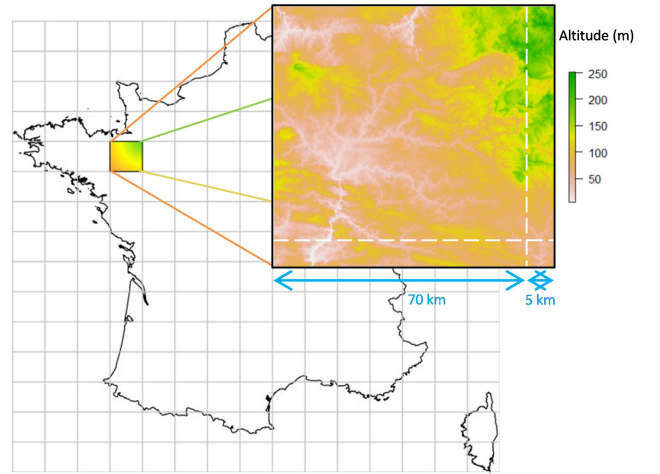


Figure 2: One raster as built in the developed method, with 5 kilometres buffer

iciency of the Hollister et al. [27] method, we applied it to 37 French lakes with previously estimated volumes<sup>3</sup>, multiplying the estimation by 0.65, in accordance with the method described in Hollister et al. [27]. Results are presented on Figure 3 and the correlation is considered sufficient. This method eliminates lakes with an area of less than about 8000 m<sup>2</sup> because of the DEM resolution. A reservoir of this area can be considered as too small for pumped storage application and does not represent a major loss of information in the evaluation process. The environmental protection zones considered

<sup>3</sup><http://adour-garonne.eaufrance.fr/>



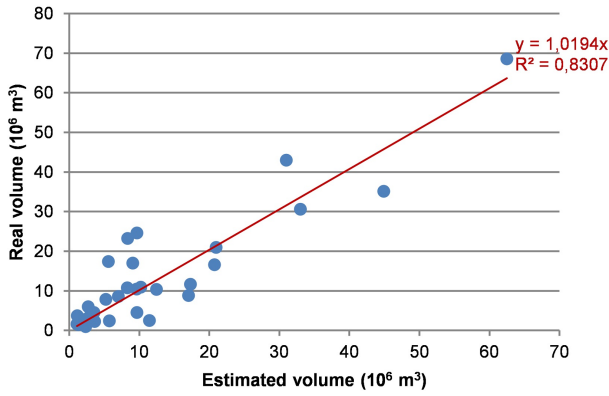


Figure 3: Results of the estimation of lake volumes on a sample of lakes with known volumes.

for this study are defined as follows, and their constraint ranking is presented in Table 5.

- **Conservatory:** Lands acquired by the national coastal and lakeside conservation authority, established to protect coastal areas and sensitive green spaces.
- **Biotope protection:** Natural zones harbouring some wild/protected animals or plants.
- **Natura 2000:** Protected because it shelters a significant share of the population of a vulnerable species. Different status level, not accessible in the database.
- **National and regional natural parks:** No fauna and flora preservation rules, but observation and control of activities in the area.
- **National and regional reserves:** Territory very strictly preserved, generally ruling out human activity.
- **International protected site:** Highly sensitive or remarkable areas protected by international institutions, like Ramsar or Unesco.

Table 5: Ranking of the different environmental and land-use constraints

Restriction	Ranking
Conservatory	No-go
Biotope protection	No-go
Natura 2000	Suitable
Nat. & Reg. parks	Suitable
Nat. & Reg. reserves	No-go
International protected site	No-go

Natura 2000 is not considered as No-go despite its environmental protection status. Such areas do not in themselves exclude PHES plants and a guide to promote PHES project development within Natura 2000 zones has even been published by stoRE [35].

In order to control the presence of buildings in natural depressions, the BATI layer from the IGN TOPO database is used. This contains all constructions higher than 20 m<sup>2</sup>, including houses, public authorities, religious buildings and sports facilities. The intersection is checked between each detected depression and building. The electric grid is also imported from the IGN TOPO database. It contains high voltage lines, from 400kV to 63kV. Employing the databases mentioned above, about 850,000 connections are detected throughout France.

For the optimization, the knapsack problem could be solved using a Mixed Integer Linear Programming solver (MILP) such as CPLEX [36] for example, and the solution can be estimated using an alternative solution. We encoded the Primal Effective Capacity Heuristic (PECH) algorithm of Akay et al. [37]. This second alternative is selected for two main reasons:

- It gives similar results compared to the exact calculation, as described in Akay et al. [37]. For the expected size of the inter-connected groups (hundreds of connections between tens of lakes maximum), this algorithm has been proven to be over 90% accurate, and possibly 95 to 99%.
- It allows for much faster computing of the optimization, and this is its biggest advantage. Indeed, the total number of potential connections can be very high: avoiding having to calculate the benefit of every combination saves considerable time.

Finally, the benefit function for the optimization process has been defined as the Energy. Indeed, we do not possess robust information to deduce an optimization of costs. The costs of making a natural depression watertight can represent a major share of the total expenditure if the soil is unsuitable, while it can be almost zero otherwise. Unfortunately, no such cost estimation has been considered possible at a large scale. A cost study lacking this information does not allow us to compare type 1, 2 and 3 connections and would imply an erroneous selection of pairings.

### 3.2. Overall potential results

After selecting the potential connections using the default restriction values (Table 4), the optimization tool gives us the results presented in Table 6. A France-wide evaluation obtains a result of about 33 GWh. The optimization step, created to ensure that every lake is related to a single connection, divides the potential by 2, confirming the capacity of the algorithm to achieve an accurate potential evaluation. The variant of the method proposed to favour the selection of pairs including existing lakes (building Type 1 and 2 first) indeed increases their proportion in the total number of connections but slightly reduces the potential. Moreover, when counting only connections involving at least one existing lake - with lower financial costs - total energy is about 22 GWh. Finally, only linking existing lakes represents a potential of about 14 GWh.

According to a report by Eurelectric [38], France possesses 184 GWh of pumped storage capacity, employing 9 plants. No

Table 6: Results of the different optimization configurations for the default restrictions. For proportions: Pink: Type 1, Gold: Type 2, Green: Type 3

Connection Nature	Prioritization	Energy before optim. (GWh)	Final Energy (GWh)	Proportions of built connections
All	No	78.42	33.56	
All	Yes	78.42	32.93	
T1 and T2	Yes	51.18	21.76	
T1	/	30.35	14.35	

new major project has been carried out since this report. The assessed potential represents an energy increase of about 18% considering all connection types, and 8% using only T1. Figure 4 presents the geographical distribution of the optimized connections across France, in terms of energy. It also gives the proportions of this energy per type of connection, as defined in Table 2. Two main observations can be made regarding this map:

- Most of the storage capacity is present in four mountainous regions: The Pyrenees, the Central Mountain Range, the Jura and the Alps.
- As shown in the pie charts, some areas offer more T1 pairing than others. In the Pyrenees (south-west France), a significant share of the total potential is provided by Type 1 connections. On the contrary, most of the energy available in the Central Mountain Range (central France) and the Jura (north-east France) is required to flood natural depressions.

Since type 1 connections are easier to build, technically and financially, we show that significant potential is concentrated in south-west France. The effects of the different steps of the evaluation process on the number of reservoirs, connections and total energy is presented in Figure 5. The cumulative dis-

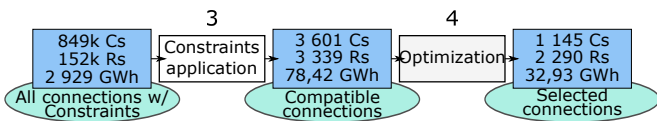


Figure 5: Evolution of number of connections and energy available in steps 3 and 4 (see Fig. 1). Cs: Connections; Rs: Reservoirs

tribution of the energy in connections is represented in Figure 6. It represents the share of the global potential achieved when considering only the connections with more energy. A logarithmic shape is observed (red dotted line, standard error: 0.003132). 80% of the potential is reached by building only 425 connections (37% of the total number).

### 3.3. Sensitivity analysis results

Five zones have been considered sensitive as defined in 2.5.1, and the variation of the estimated potential around the default acceptability constraints rankings (see Tables 3 and 5) is presented in Figure 7. Several connections are included in Natura 2000 areas and national and regional parks, which are restricted human activity zones. Some of the detected

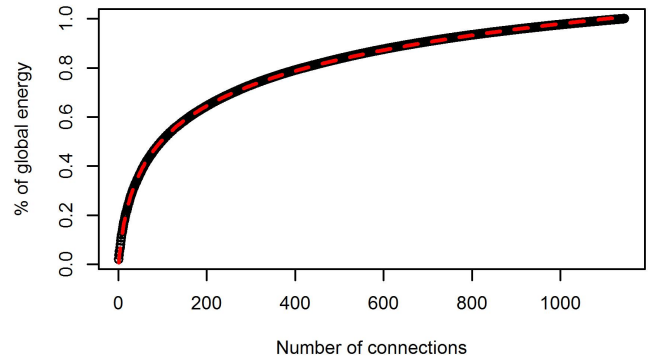


Figure 6: Cumulative distribution of energy in connections.

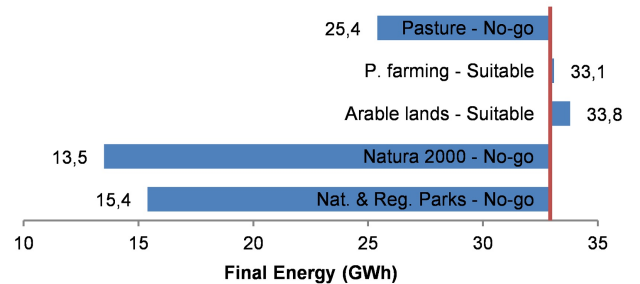


Figure 7: Variation of the evaluated storage capacity depending on some exclusion zones' status.

spots might face environmental challenges. We undertook our Global Sensitivity Analysis using 2 matrices of 400 samples of the 9 inputs described in Table 4, corresponding to 4000 evaluations of the potential. The Sobol indices estimated for two cases (including all connections and only T1) are displayed in Table 7. Estimation errors are detected, such as negative values and totals differing from 1. Both can be explained by the fact that first, the number of simulation is limited due to the time-consuming evaluation process (see 3.5), and second, we calculated only the first order of the Sobol indices. In doing so, we do not consider the potential impact of the inputs interactions in the output variability. Homma and Saltelli [34] introduced new indices expressing the total effect of a variable  $X_I$ , including the fraction of variance accounted for by  $X_I$  alone and the fraction accounted for by its combination with any other variable. These calculation of these total indices has been expressed by Jacques et al. [39], and the results for our study case are presented in Table 8.

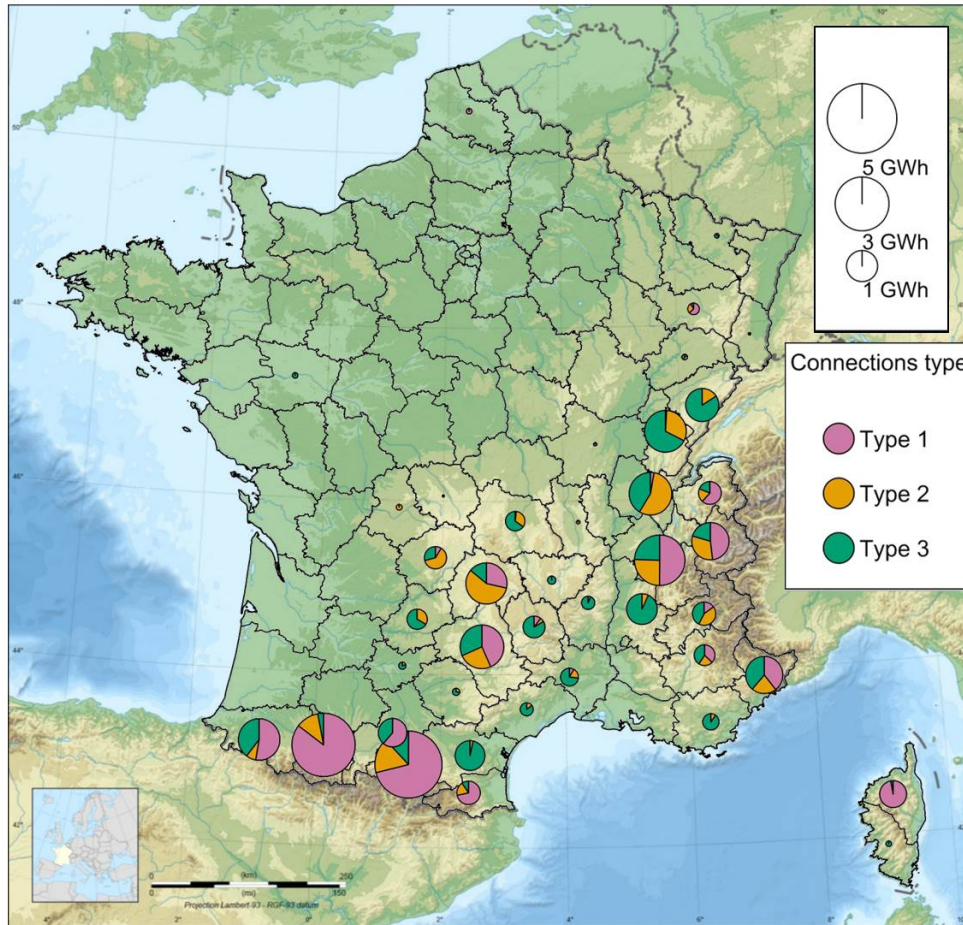


Figure 4: Distribution map of feasible connections in France, and proportion of types per county (Source of background map: <http://www.cartesfrance.fr/>)

Table 7: Sobol indices estimation for different evaluation scenarios: Case 1: Any kind of connection, with prioritization ; Case 2: Only existing lakes

Parameter	Case 1	Case 2
Min. Head	-0.003	-0.006
Max. Distance	0.285	0.119
Min. Volume	-0.012	-0.067
Max. Volume	0.015	-0.013
Min. Depth	0.023	-0.067
Max. Distance/grid	0.005	0.012
Max. Altitude	0.326	0.544
Min. Energy	0.017	-0.059
Min. H/D ratio	0.514	0.083
<b>Total</b>	<b>1.17</b>	<b>0.423</b>

Information can be extracted from the analysis. For the overall potential - including T1, T2 and T3 connections - two parameters present a high influence on the global potential: maximum distance between lakes and minimum head-distance ratio. They have a direct, and genuine influence on the evaluated potential, although their impact increases when combined with other variables. Maximum altitude has significant direct impact on the output, but this is not much greater when com-

Table 8: Sobol total indices estimation for different evaluation scenarios: Case 1: Any kind of connection, with prioritization ; Case 2: Only existing lakes

Parameter	Case 1	Case 2
Min. Head	0.391	0.111
Max. Distance	0.678	0.375
Min. Volume	0.394	0.114
Max. Volume	0.465	0.321
Min. Depth	0.411	0.118
Max. Distance/grid	0.418	0.115
Max. Altitude	0.467	0.784
Min. Energy	0.362	0.104
Min. H/D ratio	0.647	0.114

bined with other parameters.

Concerning the second case focusing on T1 connections, maximum altitude and maximum distance between lakes are the two most impacting parameters both directly and combined. Maximum volume seems to have a relative part if combined with other inputs.

### 3.4. Reasonable variations of the evaluation output

Considering the most influential factors detected above, three abacuses have been drawn representing the variation of the potential depending on these parameters, shown in Figures 8 to 10. We make them vary within a range of reasonable values, while less impacting parameters are set to their default value (Table 4). Similarly, Figure 11 represents the abacus

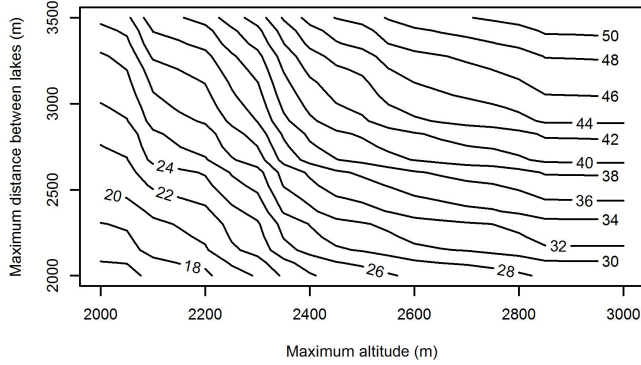


Figure 8: Energy detected (GWh) depending on the maximum distance between lakes and the maximum altitude.

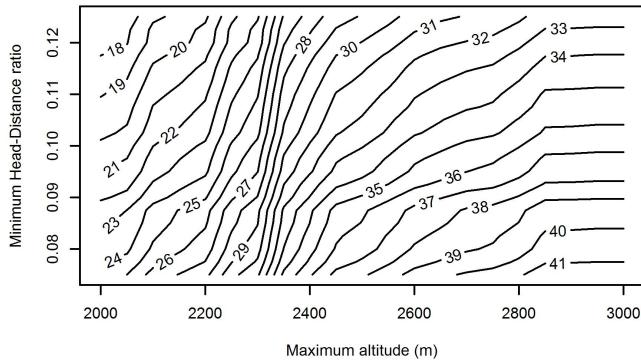


Figure 9: Energy detected (GWh) depending on the minimum head-distance ratio and the maximum altitude.

for T1 connections. Maximum altitude and distance to the grid vary in reasonable ranges, while other parameters are at their default state (Table 4). The vertical fracture visible on Figures 9 and 11 at about 2500 m altitude reveals the fast increase of the potential when accepting connection above 2300 m. Several natural lakes must be present at these altitudes. Oppositely, the impact is limited above 2500 m.

### 3.5. Computational performance

The method developed was encoded and run under RStudio [40] software, using R [41]. The simulations were performed on a 3.50 GHz desktop computer with 16 GB of RAM. This work aims to permit rapid evaluation of a large zone of energy capacity. It was thus built so that all the cumbersome and time-consuming tasks are run in one go. The actual potential evaluation steps, which require less calculation power, are run in the evaluation loop. The time-processes of the different

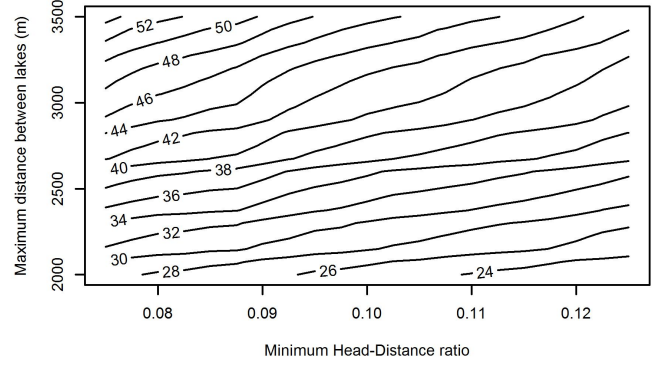


Figure 10: Energy detected (GWh) depending on maximum distance between lakes and minimum head-distance ratio.

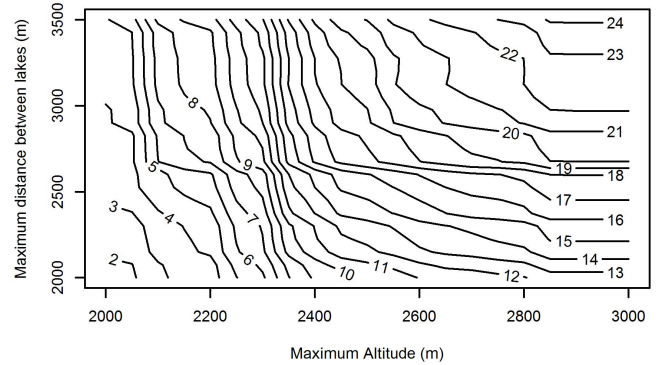


Figure 11: Energy detected depending on maximum distance between lakes and maximum altitude. Only Type 1 connections are considered.

steps are detailed in Table 9. The selection of the compatible lakes, applying environmental, land-use and technical constraints, only takes a few seconds (less than 2 seconds) while most of the time required is for the optimization process. In the simulation loop, the calculation time varies depending on the restrictiveness of the inputs, which in turn result in discarding more or less connections. The observed simulation times varied between 30 and 3,500 seconds, and a correlation curve can be drawn as on Figure 12.

## 4. Best site selection approach

Ahmadi and Shamsai [16] presented a method based on GIS data layers to rank different potential location for building an upper reservoir. Criteria are defined and each location receives a grade (ranging from 0 to 100). This makes it possible to choose the most favourable spots to be evaluated in more precise manner. This section presents a similar approach, ranking connections to select the best sites for building a small-PHES.

Part of the dataset of detected connections are considered and ranked to find which connections should be built first. A French county is considered as a demonstration case, its characteristics are detailed in Table 10.

Each connection is ranked from 0 to 100 in line with 3 criteria defined as follows, using a linear model.

Table 9: Process times of the different steps of the method.

Step	Process time	Principal difficulties
Reservoirs detection	10-15 days	Data import, Volume evaluation, Sink detection
Connections detection	3-5 days	Matching of reservoirs, Fusion of all rasters, Verification of acceptability constraints
Constraints application	2 seconds	Binary vectors selection
Optimization	30-3 500 seconds	Optimization within each inter-connected lakes group

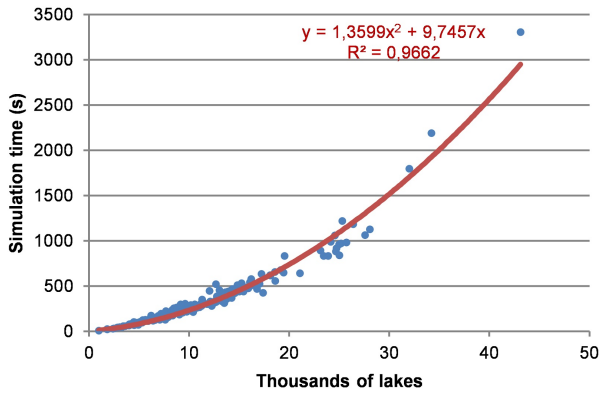


Figure 12: Evolution of process-time depending of the number of selected lakes (i.e. not restricted by constraints)

- Surface of ground to be sealed ( $R_{Surf}$ ): If a natural depression is to be filled, the cost is assumed as proportional to the volume. It is doubled if two natural depressions are paired.
- Length of the conveyance ( $R_{Dist}$ ): The greater the distance between reservoirs, the more expensive the construction of the penstock will be. The length is calculated as follows:

$$L = \sqrt{H^2 + D^2} \quad (14)$$

- Distance to electrical grid ( $R_{Dist.grid}$ ): Likewise, the greater the distance between the connection and the electrical grid, the more expensive the connection to the network will be.

Each connection receives a corresponding grade for costs ( $R_C$ ), defined as follows:

$$R_C = aR_{Surf} + \alpha(1 - a)R_{Dist} + (1 - \alpha)(1 - a)R_{Dist.grid} \quad (15)$$

$0 \leq a < 1$ : weight of the depression sealing in the cost-like function

Table 10: Characteristics of the demonstration case county.

Name	Hautes-Pyrenees
Area	4 446 km <sup>2</sup>
Number of connections detected	59
Number of T1 connection	35
Number of T2 connections	16
Total evaluated energy	4.15 GWh

$0 \leq \alpha < 1$ : balance between pipe and electrical connections weights.

The  $a$  factor represents the cost of sealing a natural depression, and clearly depends on the nature of the terrain. In a restricted area, this cost can be considered as constant for all of the connections, but if precise and more accurate data exist concerning the terrains nature and associated sealing costs,  $a$  can be used to weight the cost-like function.

Likewise, the connection energy is ranked from 1 to 100, with 100 being attributed to the highest energy value ( $R_E$ ). Finally, the "cost of energy"-like value is obtained for each connection.

$$COE = \frac{R_C}{R_E} \quad (16)$$

The results are displayed in Table 11. These results lead to two

Table 11: Results of the "Cost of energy"-like approach.

$a$	$\alpha$	ID	3 best connections	
			Type	Energy (MWh)
0	0.25	1 - 2 - 3	T2 - T1 - T1	191 - 323 - 299
	0.5	2 - 4 - 5	T1 - T1 - T1	323 - 471 - 127
	0.75	2 - 6 - 7	T1 - T1 - T1	323 - 99 - 48
>0.25	0.25	2 - 3 - 4	T1 - T1 - T1	323 - 299 - 471
	0.5	2 - 4 - 5	T1 - T1 - T1	323 - 471 - 127
	0.75	2 - 6 - 7	T1 - T1 - T1	323 - 99 - 48


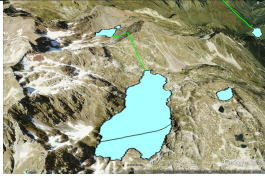
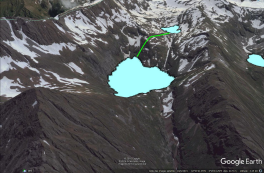
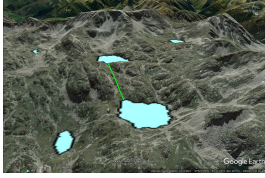



major remarks. First, Type 1 connections are widely present in the results. This can be explained by their lower cost-like ranking (no surface to be sealed) and their generally higher energy. Nonetheless, in cases where the ground does not need to be sealed, Type 2 connections can be competitive. Secondly, seven connections seemed more worthwhile than others. These connections are represented on Figure 12. Thanks to this analysis, we have been able to detect which connections might be economically advantageous. This can be of great help for a local authority keen to start implementing small-PHES on its territory.

## 5. Conclusions and perspectives

### 5.1. Conclusions

The aim of this work was to build a method able to evaluate the small-PHES potential in a massive study zone. The proposed algorithm makes it possible to assess the energy that could be stored using existing lakes or small depressions on a

Table 12: Visualization of the connections selected by the COE method. Source of background maps: Google Earth, accessed on Mar 03, 2017

ID	Visualisation	ID	Visualisation
1		5	
2		6	
3		7	
4			

terrain that could be filled. Both types of reservoirs play an important role in the evaluation given that their respective shares, taken separately, mount to less than half of the total estimated potential. Moreover, the proposed method permits fast calculation on the scale of the whole country, thanks to a global dataset generation that only needs running once. Default restrictions are suggested (e.g.. volume, depth, head, etc.) to run this analysis on any area. Applying the method to an entire country, in this case France, allowed us to affirm that the potential increases to about 33 GWh, which is almost 20% of the current national energy storage capacity. Connections are geographically divided between the Pyrenees, Central Mountain Range, Alps and Jura. The proportions of the three pairing types considered (linking existing lakes or natural terrain depressions) vary significantly within each region; composed mostly of existing lakes in the Pyrenees for example, while the potential in the Jura is based on natural depressions. Considering existing lakes only, the potential is estimated at about 14 GWh (about 8% of French current hydro storage capacity), and should be easy to mobilize. These orders of magnitude are rather high, and so we conclude that this avenue merits further study in the future.

This paper also presents a sensitivity analysis, in the form of a Sobol study. The most influential inputs are the maximum distance between lakes, their maximum altitude, and the maximum distance to the electrical grid. The corresponding abacuses are drawn to represent the output variations. Global potential varies between 16 and 54 GWh, depending on the

restrictiveness of the selected inputs.

However, the different limitations - especially the lack of economic studies and concrete sensitive zone restrictions - imply that this potential may vary in a reasonable range.

This methodology is particularly suitable for large-scale studies. It is important to note that the same process can be carried out for the case of an other country or region. The minimum necessary datasets are a DEM (SRTM is provided by NASA worldwide) and a natural waterbody database. Estimation is improved when data on environmental protection and land use are available.

The resulting connection database could also be used by hydro-power plant developers or local decision-makers to prospect for new sites to install facilities. With this mind, the last part of the study presents a different perspective. The best locations in a French county are sought, using a method similar to a "cost of energy" estimation. Some connections clearly stand out and present favourable conditions for small-PHES implementation. Moreover, the visualisation tends to confirm the effectiveness of the method, which produces coherent pairings.

## 5.2. Possible future ameliorations

Further work could be carried out on this topic. Indeed, even though the value obtained for the global French potential is probably close to reality, some details could further increase accuracy.

Connections chosen by the optimization tool within a group of inter-connected lakes are subject to discussion. The "energy" criterion is used to select the pairs, without any cost dimension. A precise cost function would eliminate the weakest connections and identify the most profitable combinations among connection groups, independently of their type. The prioritization tool would thus be obsolete.

Lake usage is not defined in the database used for this study. Some lakes used for leisure purposes were excluded using the CLC database, but it is probable that a number of existing lakes are used for other purposes, or are private, and therefore cannot be considered for PHES purposes. If a database of this kind were to be developed, a share of non-feasible reservoirs could be excluded. Moreover, due to the method used to estimate the volume of natural lakes (i.e. from surrounding topography), ponds are incorrectly evaluated. The only mean to ascertain their volume is to use real-life measurements.

The study does not include roads and transport infrastructure. Verifying the natural depression intersection with such constraints would discard other potential reservoirs. This data was too heavy to consider in this analysis, due to the very high amount of roads on the territory.

The best site selection approach may be improved, if reliable geology data are available. Each connection sealing cost could then be weighted depending on the nature of the terrain. It is likely that the global cost grade could be parametrized with cost dependence to the slope of the pipe, the distance to a road, etc.

Storage using the sea or a river as the lower reservoir was not considered. Using the sea would involve a different technology which is still in development and thus more expensive; river storage would entail difficulties in processing the river data we possess, such as river size and water flow, to ensure that they can be used as a source of water. Their inclusion in the evaluation could increase the global potential.

## Acknowledgement

The research work in this paper was carried out in the frame of the research and demonstration project "Micro-step de BERRIEN" funded in part by the French Energy Agency ADEME under the program "Energies Decarbonnées" from "Investissements d'Avenir".

## References

- [1] REN21, Renewables 2016 Global Status Report, techreport ISBN 978-3-9818107-0-7, (Paris: REN21 Secretariat), URL <http://www.ren21.net/status-of-renewables/global-status-report/>, 2016.
- [2] K. Zach, H. Auer, G. Lettner, Report summarizing the current Status, Role and Costs of Energy Storage Technologies, Tech. Rep., stoRE - Facilitating energy storage to allow high penetration of intermittent renewable energy, URL [http://www.store-project.eu/documents/results/en\\_GB/report-summarizing-the-current-status-role-and-costs-of-energy-storage-technologies/](http://www.store-project.eu/documents/results/en_GB/report-summarizing-the-current-status-role-and-costs-of-energy-storage-technologies/), 2012.
- [3] S. I. Yannopoulos, G. Lyberatos, N. Theodossiou, W. Li, M. Valipour, A. Tamburrino, A. N. Angelakis, Evolution of Water Lifting Devices (Pumps) over the Centuries Worldwide, *Water* 7 (9) (2015) 5031–5060, ISSN 2073-4441, URL <http://www.mdpi.com/2073-4441/7/9/5031>.
- [4] S. Rehman, L. M. Al-Hadhrani, M. M. Alam, Pumped hydro energy storage system: A technological review, *Renewable and Sustainable Energy Reviews* 44 (2015) 586 – 598, ISSN 1364-0321, URL <http://www.sciencedirect.com/science/article/pii/S1364032115000106>.
- [5] P. Uniyal, Old concept, new opportunities. Pump-storage, *International Water Power & Dam Construction* (2010) 20–21 URL <http://www.waterpowermagazine.com/features/featureold-concept-new-opportunities/>.
- [6] M. Manwaring, D. Mursch, K. Tilford, Challenges and Opportunities For New Pumped Storage Development, Tech. Rep., NHA s Pumped Storage Development Council, URL [http://www.hydro.org/wp-content/uploads/2014/01/NHA\\_PumpedStorage\\_071212b12.pdf](http://www.hydro.org/wp-content/uploads/2014/01/NHA_PumpedStorage_071212b12.pdf), 2012.
- [7] M. Gimeno-Gutierrez, R. Lacal-Arntegui, Assessment of the European potential for pumped hydropower energy storage based on two existing reservoirs, *Renewable Energy* 75 (2015) 856 – 868, ISSN 0960-1481, URL <http://www.sciencedirect.com/science/article/pii/S096014811400706X>.
- [8] J. Zhang, L. Xu, X. Li, Review on the externalities of hydropower: A comparison between large and small hydropower projects in Tibet based on the CO2 equivalent, *Renewable and Sustainable Energy Reviews* 50 (2015) 176 – 185, ISSN 1364-0321, URL <http://www.sciencedirect.com/science/article/pii/S1364032115004207>.
- [9] A. Nourai, Installation of the First Distributed Energy Storage System (DESS) at American Electric Power (AEP), Tech. Rep., SANDIA, 2007.
- [10] P. Denholm, E. Ela, B. Kirby, M. Milligan, The Role of Energy Storage with Renewable Electricity Generation, Tech. Rep. NREL Report No. TP-6A2-47187, National Renewable Energy Laboratory, URL [http://digitalscholarship.unlv.edu/renew\\_pubs/5](http://digitalscholarship.unlv.edu/renew_pubs/5), 2010.
- [11] S. Gabathuler, D. Pavanello, C. Mnch, Le pompage-turbinage a petite echelle pour le stockage local d energie, *Bulletin ElectroSuisse* 2/2015 (2015) 49–54 URL <https://www.hevs.ch/media/document/0/le-pompage-turbinage-a-petite-echelle-une-solution-envi.pdf>.
- [12] G. de Oliveira e Silva, P. Hendrick, Pumped hydro energy storage in buildings, *Applied Energy* 179 (2016) 1242 – 1250, ISSN 0306-2619, URL <http://www.sciencedirect.com/science/article/pii/S0306261916309874>.
- [13] N. Crettenand, The Facilitation of Mini and Small Hydropower in Switzerland, Ph.D. thesis, CDM, Lausanne, URL <https://infoscience.epfl.ch/record/176337>, 2012.
- [14] D. G. Larentis, W. Collischonn, F. Olivera, C. E. Tucci, Gis-based procedures for hydropower potential spotting, *Energy* 35 (10) (2010) 4237 – 4243, ISSN 0360-5442, URL <http://www.sciencedirect.com/science/article/pii/S0360544210003786>.
- [15] B. Kusre, D. Baruah, P. Bordoloi, S. Patra, Assessment of hydropower potential using {GIS} and hydrological modeling technique in Kopili River basin in Assam (India), *Applied Energy* 87 (1) (2010) 298 – 309, ISSN 0306-2619, URL <http://www.sciencedirect.com/science/article/pii/S0306261909003109>.
- [16] H. Ahmadi, A. Shamsai, Preliminary Site Selection of Pumped Storage Hydropower Plants - A GIS-based approach, *Energy* 35 (2009) 4237–4243, URL [http://misc.j.aut.ac.ir/article\\_237.html](http://misc.j.aut.ac.ir/article_237.html).
- [17] B. Mailler, P. Heller, N. Crettenand, Pumped-storage with small hydropower - an option?, in: SHF Pumped storage power plants conference. Lyon, France., URL <https://infoscience.epfl.ch/record/170415?ln=fr>, 2011.
- [18] S. Kucukali, Finding the most suitable existing hydropower reservoirs for the development of pumped-storage schemes: An integrated approach, *Renewable and Sustainable Energy Reviews* 37 (2014) 502 – 508, ISSN 1364-0321, URL <http://www.sciencedirect.com/science/article/pii/S1364032114003773>.
- [19] C. C. Garcia, Testing a GIS-based methodology for optimal location of Pumped Storage power plants in Norway, Ph.D. thesis, Norwegian University of Science and Technology, URL <http://hdl.handle.net/11250/242424>, 2013.
- [20] D. Connolly, S. MacLaughlin, M. Leahy, Development of a computer program to locate potential sites for pumped hydroelectric energy storage, *Energy* 35 (1) (2010) 375 – 381, ISSN 0360-5442, URL <http://www.sciencedirect.com/science/article/pii/S036054420900437X>.
- [21] R. L. Douglas Hall, Assessment of Opportunities for New United States Pumped Storage Hydroelectric Plants Using Existing Water Features as Auxiliary Reservoirs, Tech. Rep., Idaho National Laboratory, US, URL <https://indigitallibrary.inl.gov/sites/sti/sti/5998118.pdf>, 2014.
- [22] N. Fitzgerald, R. L. Arantegui, E. McKeogh, P. Leahy, A GIS-based model to calculate the potential for transforming conventional hydropower schemes and non-hydro reservoirs to pumped hydropower schemes, *Energy* 41 (1) (2012) 483 – 490, ISSN 0360-5442, URL <http://www.sciencedirect.com/science/article/pii/S0360544212001399>.
- [23] X. Lu, S. Wang, A GIS-based assessment of Tibet's potential for pumped hydropower energy storage, *Renewable and Sustainable Energy Reviews* 69 (2017) 1045 – 1054, ISSN 1364-0321, URL <http://www.sciencedirect.com/science/article/pii/S1364032116305809>.
- [24] M. Pauwels, A. Albanel, P. Blanc, F. Neirac, S. Biscaglia, Assessment of the potential of Pumped Hydro Energy Storage, in: Hydroenergia 2012 conference. Wroclaw, Poland., 2012.
- [25] J. Liebe, N. van de Giesen, M. Andreini, Estimation of small reservoir storage capacities in a semi-arid environment: A case study in the Upper East Region of Ghana, *Physics and Chemistry of the Earth, Parts A/B/C* 30 (2005) 448 – 454, ISSN 1474-7065, URL <http://www.sciencedirect.com/science/article/pii/S1474706505000409>.
- [26] T. Sawunyama, A. Senzanje, A. Mhizha, Estimation of small reservoir storage capacities in Limpopo River Basin using geographical information systems (GIS) and remotely sensed surface

- areas: Case of Mzingwane catchment, *Physics and Chemistry of the Earth, Parts A/B/C* 31 (2006) 935 – 943, ISSN 1474-7065, URL <http://www.sciencedirect.com/science/article/pii/S1474706506001677>.
- [27] J. W. Hollister, W. B. Milstead, M. A. Urrutia, Predicting Maximum Lake Depth from Surrounding Topography, *PLoS ONE* 1 (9) (2011) 6, URL <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0025764>.
- [28] O. Planchon, F. Darboux, A fast, simple and versatile algorithm to fill the depressions of digital elevation models, *CATENA* 46 (2 - 3) (2002) 159 – 176, ISSN 0341-8162, URL <http://www.sciencedirect.com/science/article/pii/S0341816201001643>.
- [29] S. Papaefthimiou, E. Karamanou, S. Papatthassiou, M. Papadopoulos, Operating policies for wind-pumped storage hybrid power stations in island grids, *IET Renewable Power Generation* 3 (2008) 293–307, URL <http://ieeexplore.ieee.org/document/5237666/?arnumber=5237666>.
- [30] S. A. Berezinskii, A. V. Egorov, V. S. Lashmanova, I. A. Pollnkovskii, Design of penstocks of pumped-storage stations, *Hydrotechnical Construction* 4 (1985) 180–187, URL <https://link.springer.com/article/10.1007%2FBF01429267>.
- [31] M. R. Garey, D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*, W. H. Freeman & Co., New York, NY, USA, ISBN 0-7167-1045-5, 1990.
- [32] P. Padey, R. Girard, D. Le Boulch, I. Blanc, From LCAs to Simplified Models: A Generic Methodology Applied to Wind Power Electricity, *Environmental Science and Technology* 47 (3) (2013) 1231–1238, URL <http://pubs.acs.org/doi/abs/10.1021/es303435e>.
- [33] I. Sobol, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Mathematics and Computers in Simulation* 55 (2001) 271 – 280, ISSN 0378-4754, URL <http://www.sciencedirect.com/science/article/pii/S0378475400002706>, the Second {IMACS} Seminar on Monte Carlo Methods.
- [34] T. Homma, A. Saltelli, Importance measures in global sensitivity analysis of nonlinear models, *Reliability Engineering & System Safety* 52 (1) (1996) 1 – 17, ISSN 0951-8320, URL <http://www.sciencedirect.com/science/article/pii/0951832096000026?via%3Dihub>.
- [35] M. K. nee Reidy, Development of Bulk Energy Storage and Natura 2000, Tech. Rep., stoRE - Facilitating energy storage to allow high penetration of intermittent renewable energy., URL [http://www.store-project.eu/documents/results/en\\_GB/development-of-bulk-energy-storage-natura-2000](http://www.store-project.eu/documents/results/en_GB/development-of-bulk-energy-storage-natura-2000), 2014.
- [36] IBM, IBM ILOG CPLEX Optimizer, URL <http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/>, 2010.
- [37] Y. Akay, H. Li, S. H. Xu, A Greedy Algorithm for the General Multidimensional Knapsack Problem, *Annals of operations research* 150, URL <https://link.springer.com/article/10.1007%2Fs10479-006-0150-4>.
- [38] Eurelectric, Hydro in Europe: Powering Renewables, Tech. Rep., Eurelectric, URL [http://www.eurelectric.org/media/26690/hydro\\_report\\_final-2011-160-0011-01-e.pdf](http://www.eurelectric.org/media/26690/hydro_report_final-2011-160-0011-01-e.pdf), 2011.
- [39] J. Jacques, C. Lavergne, N. Devictor, Sensitivity analysis in presence of model uncertainty and correlated inputs, *Reliability Engineering & System Safety* 91 (1011) (2006) 1126 – 1134, ISSN 0951-8320, URL <http://www.sciencedirect.com/science/article/pii/S0951832005002231>, the Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004) SAMO 2004 The Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004).
- [40] RStudio Team, RStudio: Integrated Development Environment for R, RStudio, Inc., Boston, MA, URL <http://www.rstudio.com/>, 2015.
- [41] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org/>, ISBN 3-900051-07-0, 2013.