

How effective are heuristic solutions for electricity planning in developing countries

Yakubu Abdul-Salam ^{a,*}, Euan Phimister ^b

^{a,*} Corresponding author: Yakubu Abdul-Salam, The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH; yakubu.abdul-salam@hutton.ac.uk; Tel: +44(0)1224395453

^b Aberdeen Centre for Research in Energy Economics and Finance, University of Aberdeen, Aberdeen, UK; e.phimister@abdn.ac.uk; Tel: +44(0)1224273855

How effective are heuristic solutions for electricity planning in developing countries.

Abstract

Heuristic algorithms have been widely used to provide computationally feasible means of exploring the cost effective balance between grid versus off grid sources for universal electrification in developing countries. By definition in such algorithms however, global optimality is not guaranteed. We present a computationally intensive but globally optimal mixed integer non-linear programming (MINLP) model for electricity planning and use it in a Monte Carlo simulation procedure to test the relative performance of a widely used heuristic algorithm due to Parshall et al. (2009). We show that the overall difference in cost is typically small suggesting that the heuristic algorithm is generally cost effective in many situations. However we find that the relative performance of the heuristic algorithm deteriorates with increasing degree of spatial dispersion of unelectrified settlements, as well as increasing spatial remoteness of the settlements from the grid network, suggesting that the effectiveness of the heuristic algorithm is context specific. Further, we find that allocation of off grid sources in the heuristic algorithm solution is often significantly greater than in the MINLP model suggesting that heuristic methods can overstate the role of off-grid solutions in certain situations.

Keywords: Electricity, Algorithms, Mixed integer programming, Grid/Off-grid, Monte Carlo simulation, Parshall et al algorithm

1 Introduction

It is estimated that up to 1.3 billion of the world's population have no access to electricity and a majority reside in rural areas of the world's developing regions (IEA WEO, 2013). The situation is most pronounced in sub Saharan Africa where the overall electrification rate is about 30% only, with some countries such as Chad, Liberia and South Sudan having less than 5% electricity coverage (World Bank, 2015).¹ There is a positive correlation between access to electricity and development (Goldemberg, 2000) and although access to electricity in itself is not a panacea for development (Bhattacharyya, 2006), modest access to it can have significant impact on the general wellbeing² of the poor (Khandker et al., 2009). In many cases however, the poor in developing countries live in rural locations that are thinly inhabited and remote from existing national grids so that high fixed costs of grid extension means grid access to these locations is often uneconomical. Meanwhile potential for the use of off-grid technologies in these countries, particularly drawing on renewable resources such as wind, solar and (bio)diesel has been promoted by a range of authors including Buys et al. (2007) and Painuly and Fenhann (2002).

A range of algorithms focussing on cost effective planning methods for universal electrification, incorporating both the economic and networking aspects of grid and off-grid electricity planning in developing countries have been proposed. Among these are the algorithms by Lambert and Hittle (2000), Amador and Dominguez (2005), Parshall et al. (2009), Deichmann et al. (2011), Levin and Thomas (2012), Szabó et al. (2013) and Abdul-Salam and Phimister (2016). Owing to the complexity and large scale nature of the underlying optimization problem,³ mathematical models based on combinatorial optimisation techniques are impractical. These algorithms are therefore based on heuristic methods. Being ad-hoc heuristics however, there is by definition no measure of the degree to which they are cost effective and they may in fact provide very different solutions not just in terms of cost but also in terms of the spatial frontiers of the competing grid and off-grid technologies (Abdul-Salam and Phimister, 2016).

In this paper, we present a mixed integer non-linear programming (MINLP) formulation of the underlying cost minimization problem and use it to test the relative performance of the heuristic algorithm by Parshall et al. (2009) (herein referred to as the PA algorithm). We choose the PA algorithm for two reasons. First, it has been widely adopted in the literature and its advanced user interface program is freely available online.⁴ The algorithm has been

¹ Inadequate maintenance of ageing grid network infrastructure as well as insufficient generation capacities among other financial, technical and management issues means that even for electrified locations, service is often unreliable.

² For example, electricity used for lighting can decrease the incidence of respiratory diseases which result from use of smoke producing biomass based indoor lanterns.

³ Parshall et al. (2009) modelled 6612 nodes, Deichmann et al. (2011) modelled between 700 and 1000 nodes and Abdul-Salam and Phimister (2016) modelled 1086 nodes.

⁴ The PA algorithm is accessible at <http://networkplanner.modilabs.org/docs/>. We however implement both the MINLP model and PA algorithm in GAMS. The codes are available upon request.

used to study electricity planning in Kenya (Parshall et al., 2009), Senegal (Sanoh et al., 2012), Ghana (Abdul-Salam and Phimister, 2016; Kemausuor et al., 2014), Nigeria (Akpan, 2014; Ohiare, 2015) and gas network planning in East Africa (Demierre et al., 2015). Second, it has been tested against two existing heuristic algorithms and have been found to yield better cost effective results (see Abdul-Salam and Phimister, 2016). For these reasons, the PA algorithm can be regarded as the benchmark heuristic for cost effective electricity planning in developing countries. By definition, the MINLP model introduced in this paper yields globally optimum solutions hence provides a basis for testing the performance of the PA algorithm. However the complexity of the cost effective universal electricity planning problem renders the MINLP model to be NP complete hence it is only able to reliably solve small scale planning problems involving up to 40 settlements only.⁵ Although this limit is small, many rural electrification projects in developing countries typically involve sub-regional planning situations involving small numbers of communities considered at a time (e.g. Maristes, 2011; World Bank, 2007, World Bank, 2013) hence making the MINLP model a practically useful tool in those circumstances. We present an example sub-regional planning problem in Section 0 to show how the model might be used for such situations.

We test the relative performance of both methods by applying them to a small and simplified version of the universal electricity planning problem. To provide a more realistic setting for the simulations we use grid and off-grid electrification cost data from Ghana. Spatial factors are known to be important determinants of electrification costs (Zvoleff et al., 2009). To explore how these factors affect the relative performances of both methods, we simulate the location of the settlements which are to be electrified using Monte Carlo simulation and explore how the relative costs of the two methods vary with respect to two spatial metrics, namely degree of dispersion between the simulated settlements, and their degree of remoteness from the existing grid network.

We show that the overall average difference in performance between the two solutions is small. Across all Monte Carlo trials, the average percentage absolute error between the two solutions is only 0.7%, with the maximum of around 3.8%. This suggests that the PA algorithm is in general an effective planning tool, and that it provides a good yardstick against which to judge other heuristic algorithms in use for large scale electricity planning. However, we do find that the relative performance of the PA algorithm deteriorates with increasing degree of dispersion between unelectrified settlements, and increasing remoteness of the settlements from the existing grid and that more off-grid technology use is implied within the heuristic PA algorithm than the MINLP model.

The remainder of the paper is organised as follows. In Section 2, we describe in more detail the nature of the electricity planning problem and then outline the PA algorithm and the MINLP model developed in this paper. We also use a stylised example to show how both methods work. In Section 3, we describe the Monte Carlo simulation procedure, the data and the evaluation measures used. In Section 4, the Monte Carlo simulation results are discussed in detail. We also present results of the MINLP model applied to a sub-regional planning problem from Ghana. Section 5 concludes with policy implications.

⁵ This limit regards running the model on a standard desktop computer with a 4GB RAM, a 64 bit operating system and 2.70GHz frequency. With greater computing resource, this limit could be extended.

2 Methodology

2.1 The nature of the problem

Figure 1 provides an illustration of the spatial nature of the planning problem for Ghana, showing the existing centralised grid infrastructure and some unelectrified settlements (represented by the isolated dots). Point settlements and high voltage (HV) transmission lines are actual data for Ghana (see SWERA, 2011). The medium voltage (MV) lines are suggested grid extension frontiers shown in Abdul-Salam and Phimister (2016). Given the policy objective of achieving universal access to electricity at minimum cost, the problem posed here is to determine where and how to extend the centralised grid and where to provide off-grid solutions incorporating renewables.⁶ Clearly the potential solution to this problem will be sensitive to the relative economics of networked grid extension and the competing off-grid renewable system options, which in turn will depend on spatial resource availability. It will also be sensitive to factors such as demand for electricity in individual settlements, closeness of settlements to the existing grid, their demographic and economic characteristics (e.g. population, population density), etc.

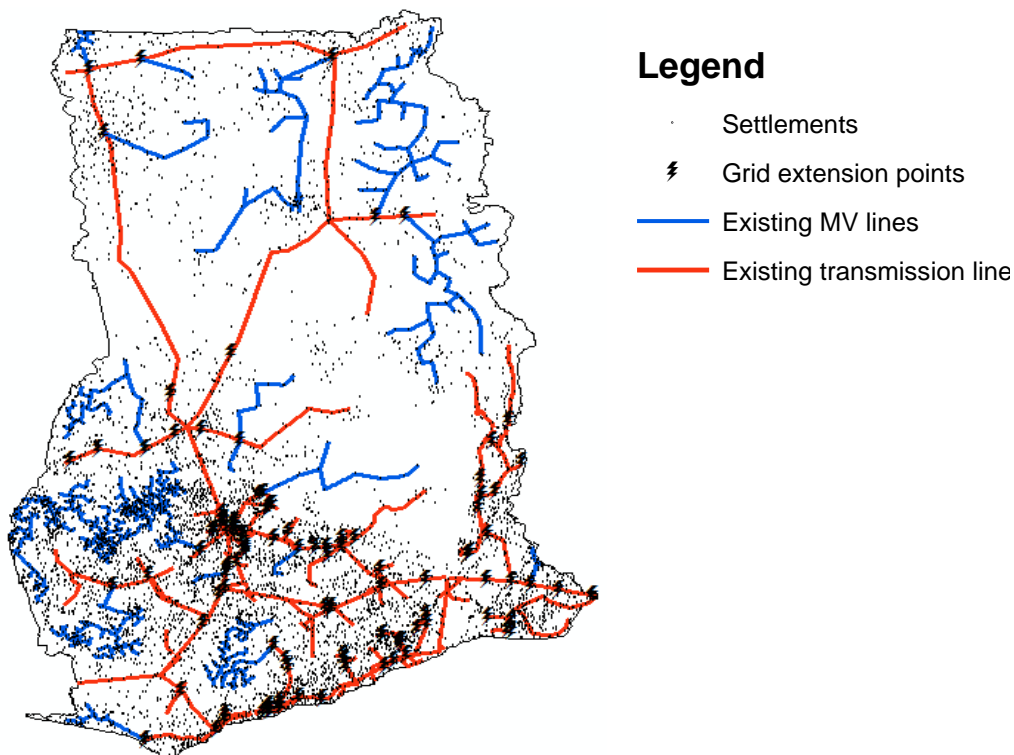


Figure 1: The cost effective universal electrification problem configured for Ghana. Source: Adapted from Abdul-Salam (2015).

⁶ The off-grid systems we consider are standalone household solar and wind systems and autonomous biodiesel mini-grid systems.

The complexity of the underlying optimization problem arises from the need to determine simultaneously the optimal balance between grid and off grid systems, with the optimal configuration of the distribution network for grid assigned nodes. Typically, a minimum spanning tree (MST) method is used to route grid assigned nodes, forming a radial electricity distribution network. However, finding the MST of a set of nodes analytically is challenging as it belongs to the class of NP-complete problems including the travelling salesman problem (Magnanti, 1995). Fast heuristics for the MST algorithm however exist (e.g. Prim, 1957; Kruskal, 1956). This has naturally led to the use of heuristics in order to deal with large scale universal electricity planning problems where thousands of unelectrified nodes need to be considered.

This type of "macro" level electricity planning problem dealt with here can be contrasted with other types of electricity planning problems previously considered in the literature dealing typically with different parts of the overall problem. Many papers have considered the detailed planning of the distribution system considering the design and optimisation of utility scale grid networks only (see Sempertegui et al. (2002) for an overview of this literature). This literature typically does not consider off-grid renewable systems such as solar and wind technologies and in situations where these are considered, their use is not examined as autonomous systems but rather they are integrated into the central grid network (i.e. distributed generation). Cost minimisation is often not the focus but rather system stability and reliability, minimisation of energy loss, etc. Network constraints, through load flow analyses, are given significant treatment in this literature. The size of the problems considered are typically small (i.e. up to tens of nodes). While this literature has typically not considered increasing access to electricity in the developing world, related work has considered the detailed design of micro-grids. For example, Ferrer-Martí et al. (2011) consider the optimal design of a microwind rural electrification systems to minimize cost. A separate strand in the literature has dealt with developing (often ad-hoc) criteria for identifying settlements for grid and/or off-grid electrification in the context of enhancing access to electricity in the developing world (e.g. World Bank, 2008; Nguyen, 2007; Ranaboldo, 2013). The planning goal is often driven by government policy which may not necessarily be based on a goal to minimise cost but rather altruistic goals such as bridging the access to electricity gap between regions, communities, etc. Such goals are consistent with the notion of 'just grids' as discussed by Bazilian et al. (2011) where increased access to energy services are designed in a manner that will not marginalise the poor.

2.2 The PA algorithm

As earlier stated, this algorithm is widely used and can be regarded as the benchmark heuristic for the cost effective universal electricity planning problem. The algorithm first computes the *internal grid cost* for all unelectrified settlements or nodes⁷ in an electricity planning problem. For each node, this cost is computed as the sum of the cost of connecting its households and institutions including the cost of secondary MV lines, MV to low voltage

⁷ Although throughout the paper we use the term settlement interchangeably with node, in the more technical sections we use the term node as it is technically more precise.

(LV) transformers, LV lines, internal household wiring costs, tariff payments for electricity service, etc. As shown in Figure 2, these grid related costs are internal to the geographic boundaries of a node hence *internal grid costs*. It does not include the cost of the new primary MV backbone that is needed to connect a node to the MV backbone of the existing grid network. New primary MV backbone lines are largely external to the geographic boundaries of a node.

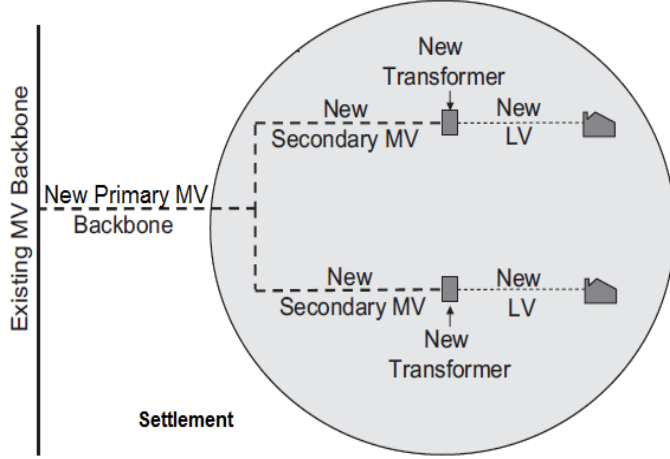


Figure 2: Defining the components of ‘internal grid cost’. Source: Parshall et al. (2009)

For each unelectrified node, the algorithm also calculates the cost of the off-grid technologies under consideration. The off-grid technologies we consider are autonomous solar, wind and diesel mini-grid systems. The following initial determinations are then made;

1. If the *internal grid cost* for an unelectrified node is higher than the cost of at least one of the off-grid technologies under consideration, the node is marked to be 'ineligible' for grid connection. Clearly grid connection is not viable if an off-grid technology is cheapest choice even before considering grid cost of building a new primary MV line. Such a node is served its least cost off-grid technology.
2. If the *internal grid cost* for an unelectrified node is less than the cost of all off-grid technologies being considered, that node is identified to be 'eligible' for grid connection and is assigned a metric called MV_{\max} .

The MV_{\max} metric for an eligible node is defined as the maximum allowable length of a new primary MV line to be extended from the existing MV backbone to that node such that the total grid cost (i.e. *internal grid cost* plus cost of incoming primary MV line) is less than or equal to the cost of the cheapest off-grid technology for the node. Consistent with this definition, we calculate MV_{\max} for an eligible node k as follows;

$$\min_{MV_{\max}(k)} npc_{cheapest_offgrid}(k) = npc_{internal_grid}(k) + unitCost_{MV} \times MV_{\max}(k) + \sum_{t=1}^T \frac{OandM_{MV} \times MV_{\max}(k)}{(1+r)^{t-1}} \quad (1)$$

where $npc_{cheapest_offgrid}(k)$ and $npc_{internal_grid}(k)$ are the net present costs of the cheapest off-grid technology at eligible node k and its *internal grid cost* respectively; T and r are the planning horizon and discount rate respectively; $unitCost_{MV}$ and $OandM_{MV}$ are the unit capital cost (\$/km) of an MV line and its unit operations and maintenance cost (\$/km) of respectively.

At each iteration, one eligible unelectrified node is connected to the existing grid network. The connected node is served with an MV extension that is less than or equal to its MV_{max} metric. These connections are based on the MST algorithm. Given a set of nodes, the MST algorithm finds the network of lines that connects all the nodes such that the total length of lines is the minimum possible, and that no loops are present. The PA algorithm terminates when all or at least one of the following conditions are reached; (1) all eligible unelectrified nodes have been connected to the network; or (2) the remaining eligible unelectrified nodes are located further from the extended grid network than their MV_{max} distance metric. In that case, they are simply assigned their cheapest off-grid technology.

2.3 The MINLP model

As previously stated, the cost effective universal electricity planning problem typically involves many settlements which makes the problem large and impractical for mathematical programming techniques. In this section, we introduce an MINLP model which we develop to handle a smaller version of the problem in order to investigate the relative performance of the heuristic algorithms used in larger problems. The version of the planning problem we consider involves multiple unelectrified nodes and a single grid source node. All unelectrified nodes that are determined to be optimal for grid electrification are to be connected to this source node to form a grid distribution network. Hence let I represent the set of all nodes in the planning problem, let $G \in I$ represent the set of all unelectrified nodes and let $s \in I$ represent the single grid source node. For a node $i \in I$, let $npc_{minigrd}(i)$, $npc_{solar}(i)$, and $npc_{wind}(i)$ represent the exogenously calculated net present cost of electrifying the node with the respective off-grid technologies i.e. mini-grid, solar and wind. Also let $npc_{internal_grid}(i)$ represent the exogenously calculated *internal grid cost* for the node. This cost is as previously defined for the case of the PA algorithm. We assume that inter-nodal distances are Euclidean and represent this distance with the exogenous parameter $distance(i, j)$ where node j is an alias for node i . Similar to the PA algorithm, we assume that all grid assigned nodes are networked via a MST algorithm to form their grid distribution network. Hence let $g \in G$ represent an unelectrified node. We define a parameter θ which controls the formation of the MST distribution network as follows;

$$\begin{aligned} \theta(s, g) &= -1 \\ \theta(g, g) &= 1 \end{aligned} \tag{2}$$

Now let $w(i), x(i), y(i)$ and $z(i)$ represent endogenously determined binary variables indicating the assignment of node i for mini-grid, solar, wind or the grid respectively. Also let $x_mst(i, j)$ represent endogenously determined binary variable showing the MST network of all grid assigned nodes into the existing grid. $length_{MST}$ is endogenous variable showing the total length of the MST network and $y_mst(i, j, g)$ is binary flow indicator variable that directs the endogenous MST network generation.

Given the above parameters and variables, we formulate the MINLP model as follows;

Cost minimisation and technology assignment equations:

$$\begin{aligned}
\min_{\substack{w(i),x(i),y(i),z(i), \\ x_mst(i,j),y_mst(i,j,g)}}} \text{cost} = & \sum_i npc_{minigrd}(i) \times w(i) + \sum_i npc_{solar}(i) \times x(i) + \sum_i npc_{wind}(i) \times y(i) \\
& + \sum_i npc_{internal_grid}(i) \times z(i) + unitCost_{MV} \times length_{MST} \\
& + \sum_{t=1}^T \frac{OandM_{MV} \times length_{MST}}{(1+r)^{t-1}}
\end{aligned} \tag{3}$$

s.t.

$$\sum_i (w(i) + x(i) + y(i) + z(i)) = N \tag{4}$$

$$w(i) + x(i) + y(i) + z(i) = 1 \quad \forall i \in I \tag{5}$$

$$z(s) = 1 \quad \forall s \in I \tag{6}$$

MST networking equations imposed as further constraint:

$$length_{MST} = \sum_{i,j} (distance(i,j) \times x_mst(i,j)) \tag{7}$$

$$\begin{aligned}
& \sum_{i,j} (y_mst(i,j,g) \times z(i) \times z(j) \times z(g)) + \sum_{j,i} (y_mst(j,i,g) \times z(j) \times z(i) \times z(g)) \\
& = \theta(i,g) \times z(i) \times z(g) \quad \forall i \in I, g \in I, g \in G
\end{aligned} \tag{8}$$

$$\sum_{i,j} (x_mst(i,j)) = \left(\sum_i z(i) \right) - 1 \tag{9}$$

$$x_mst(i,j) \geq y_mst(i,j,g) \quad \forall i \in I, j \in I, g \in I, g \in G \tag{10}$$

$$i \neq j, i \in I, j \in I, s \in I, g \in I, g \in G, s \notin G$$

N is the total number of nodes, T and r are as previously defined. θ is as defined in equation (2), Equation (3) seeks a minimisation of the total net present cost of grid and off-grid electrification. Equation (4) constrains the sum of the binary endogenous variables to equal the total number of nodes in the problem. This is to say we cannot have more assignments of technologies than there are nodes in the problem. Equation (5) constrains one technology assignment to each node; it complements equation (4). Equation (6) states that the grid generation node or primary MV connection point s is necessarily grid assigned.

Endogenous binary variables $w(i), x(i), y(i)$ and $z(i)$ indicate which technology is assigned to a node i . For a node i that is assigned an off-grid technology [i.e. $w(i)=1$ or $x(i)=1$ or $y(i)=1$], the total cost of electrification captured in exogenously calculated parameters [$npc_{minigrid}(i)$ or $npc_{solar}(i)$ or $npc_{wind}(i)$] respectively for that node suffices as its total electrification costs. For any node assigned to the existing grid however [i.e. $z(i)=1$], exogenously calculated internal grid cost $npc_{internal_grid}(i)$ is only the total grid cost that is internal to the node. There is an additional cost for grid assigned nodes. This is the cost of building the MST network connecting them together and into the existing grid. The MST network is generated by the set of constraints in equations (7)-(10). Notice that these set of constraints are in relation to grid assignment variable z only and not the off-grid assignments w, x, y . The total length of the MST network generated is captured by the variable $length_{MST}$. The cost of the MST network, which as previously noted is constructed with MV lines, is captured by the terms $unitCost_{MV} \times length_{MST}$ (i.e. capital costs) and $\sum_{t=1}^T \frac{OandM_{MV} \times length_{MV}}{(1+r)^{t-1}}$ (i.e. operations and maintenance costs over T years) in equation (3).

Simulations show that the MINLP model is computationally very intensive. For small sized problems of up to 40 nodes, the computational resource requirement is significantly higher than the PA algorithm. For larger sized problems, computer resource needs increase rapidly and solver software (SCIP in GAMS software) fail. The reason for the computational intensity of the model is the computational cost of solving the MST imposed constraint for networking nodes. As stated previously, a mathematical formulation of the MST algorithm is known to be NP complete. Table 1 demonstrates the computational intensity of the MINLP model. For comparison, the computational time of the PA algorithm are provided in parentheses in the fifth column of Table 1.

Total number of nodes	Number of single variables	Number of single equations	Number of non-zero elements	Computational time ⁸
10	949	913	6,719	15 mins (8 secs)
20	7,699	7,623	58,449	45 mins (8 secs)
30	26,249	26,133	203,179	110 mins (8 secs)
40	62,599	62,443	488,909	300 mins (8 secs)
50	122,749	122,553	963,639	Solver fail (8 secs)
100	990,301	989,905	7,850,709	Solver fail (8 secs)

Table 1: Demonstrating computational intensity of the MINLP model (comparative running times of the PA algorithm in parentheses).

⁸ Computational times from simulating both methods on a 4GB RAM computer with a 64 bit operating system and 2.70GHz frequency.

Table 1 shows that for problem sizes upto 40 nodes, the computational burden of the MINLP model is significantly higher than the heuristic PA algorithm. For sizes up to 50 nodes and over, the solver for the MINLP model fails. Due to the computational intensity of the MINLP model shown in Table 1, the model is unsuitable for large sized universal electricity planning problems of the kinds handled by the heuristic approaches mentioned. However it can be used for smaller sized versions of the problem involving few nodes e.g. at the sub-regional planning level (e.g. Maristes, 2011; World Bank, 2007, World Bank, 2013). An example sub-regional planning problem is provided in Section 0, showing how the model may be used in such cases.

2.4 A Stylised Example Planning Problem.

To demonstrate how the two methods can work, consider the highly stylised problem represented in Figure 3 and Table 2. To emphasise the potential differences between the two methods we have chosen an extreme example, where in contrast to the MINLP model solution, the PA algorithm will choose not to extend the grid.

In this example there are 8 nodes (N1-N8) representing unelectrified settlements and one node which represents the connection point to the existing grid (S1). In Table 2 the net present value of costs are given for electrification of each node using the grid and off-grid methods. The planning objective is to achieve access to electricity for all nodes at minimum total cost by determining which nodes to serve by grid (and how to extend the grid for those cases) and which nodes to serve using off-grid solutions such as renewables (e.g. autonomous solar, wind and biodiesel mini-grid systems).



Figure 3: Stylised example planning problem: Spatial distribution of unelectrified settlements.

Node	Coordinates		Distance to S1 (km)	Net present cost, \$ ('0000)			
	Latitude	Longitude		$npc_{internal_grid}$	$npc_{minigrid}$	npc_{solar}	npc_{wind}
N1	1	7	17.49	<u>50</u>	60	74	76
N2	3	10	14.32	<u>50</u>	60	74	76
N3	6	9	12.21	<u>50</u>	60	74	76
N4	5	7	14.21	<u>50</u>	60	74	76
N5	2	13	14.32	<u>50</u>	60	74	76
N6	5	25	14.21	80	<u>70</u>	72	74
N7	25	2	16.64	80	72	<u>70</u>	74
N8	25	25	12.73	80	74	72	<u>70</u>
S1	16	16	-	-	-	-	-

Table 2: Stylised example planning problem: Cost Data.

Nodes N1-N5 are highly clustered in the South West region of space in Figure 3 whilst N6-N8 are dispersed. Applying the PA algorithm, nodes N1-N5 are classified as ‘eligible’ for grid electrification because their internal grid cost $npc_{internal_grid}$ is cheapest whilst N6, N7 and N8 are classified as ‘ineligible’ because their off-grid costs are cheaper. N6, N7 and N8 are therefore assigned their cheapest off-grid technologies namely mini-grid, solar and wind technologies respectively. Applying equation (1) with the parameters planning horizon $T=10$ years, interest rate $r=10\%$ and operations and maintenance cost of MV line

Given $O_{and}M_{MV} = \$282/km$, we find the MV_{max} metric for each of the 5 eligible nodes to be 6.3 km. As none of these nodes is within 6.3 km of the grid source node S1 (see Table 2), they are ineligible for grid connection and are therefore awarded their least cost off-grid technologies i.e. the mini-grid in each case. The final solution of the PA algorithm therefore does not involve connecting any node to the grid generation source.

Unlike the PA algorithm however, applying the MINLP model to the same problem results a solution with a grid network extended from the grid source node S1 to connect the 5 clustered nodes N1-N5. The total length of this network is 24.4km as shown in Figure 3.

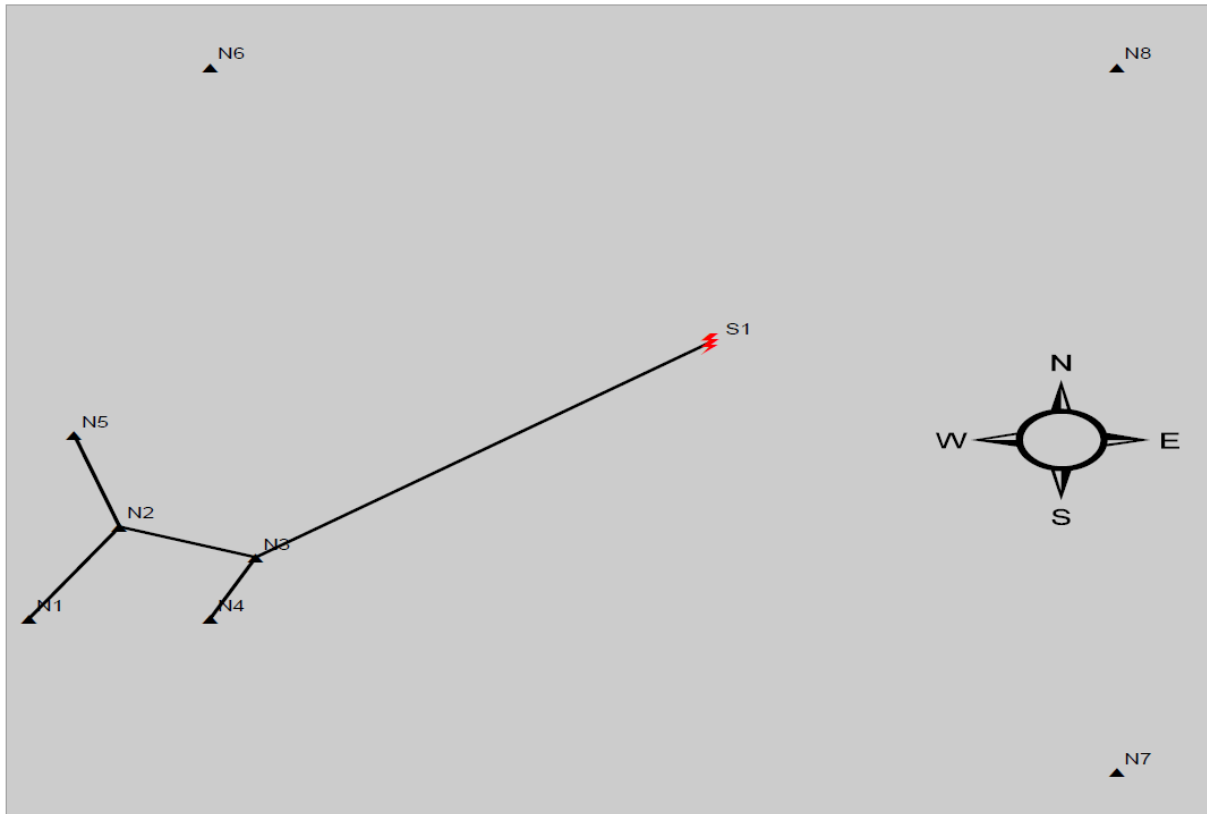


Figure 4: MINLP model solution for stylised problem

The total cost of the MINLP model and the PA algorithm are \$4.99million and \$5.10million respectively, a difference of \$110,000 over 10 years. The key to the lower cost solution in the MINLP model is its recognition of the high clustering of nodes N1-N5 in the South West region of space, and the low internal grid costs of those nodes. High clustering provides opportunities for lowering grid connection costs due to lower network wiring costs. The PA algorithm in this instant is less cost effective because the highly clustered region is sufficiently distant from the grid source S1, hence barring grid connection to the clustered nodes due to the MV_{max} criterion.

3 Monte Carlo procedure

To explore the relative performance of the two methods more robustly, we use Monte Carlo simulation to capture how differences in spatial factors and settlement electrification costs affect the relative solutions. The Monte Carlo experiment is structured as follows. In each trial (of 434), we randomly generate the position of 21 unelectrified nodes in an artificially defined space.⁹ For simplicity, only one grid source node is generated. This node represents the existing grid and is the point to which all grid assigned demand nodes are connected. To increase the likelihood of variations in spatial dispersion and remoteness, we vary the spatial extents of the spaces within which nodes are simulated. Higher spaces increase the likelihood of high dispersion and remoteness and vice versa for smaller spaces.

3.1 Cost Data

To generate realistic cost distributions for each type of electrification method, costs of grid and off-grid electrification (i.e. standalone household solar and wind systems and autonomous mini-grid systems) for each demand node are picked from distributions of grid and off-grid electrification costs in Ghana using data constructed by Abdul-Salam (2015). The study combined geospatial data for solar and wind resource potential of unelectrified settlements in Ghana (SWERA, 2011) with settlement populations, Afripop data (2011), plus assumptions about technology costs and configurations, to calculate the distribution of net present costs for over 1000 unelectrified nodes in Ghana by electrification method. The derived empirical distributions which are used to draw the net present costs for each node in the Monte Carlo simulation are illustrated in Figure 5.

⁹ The MINLP model can be solved using SCIP solver in GAMS for up to 40 nodes. However experimentation showed that the nature of the results were not affected when using a smaller number of nodes hence the choice of 21 nodes. This eased the computational burden and solution time.

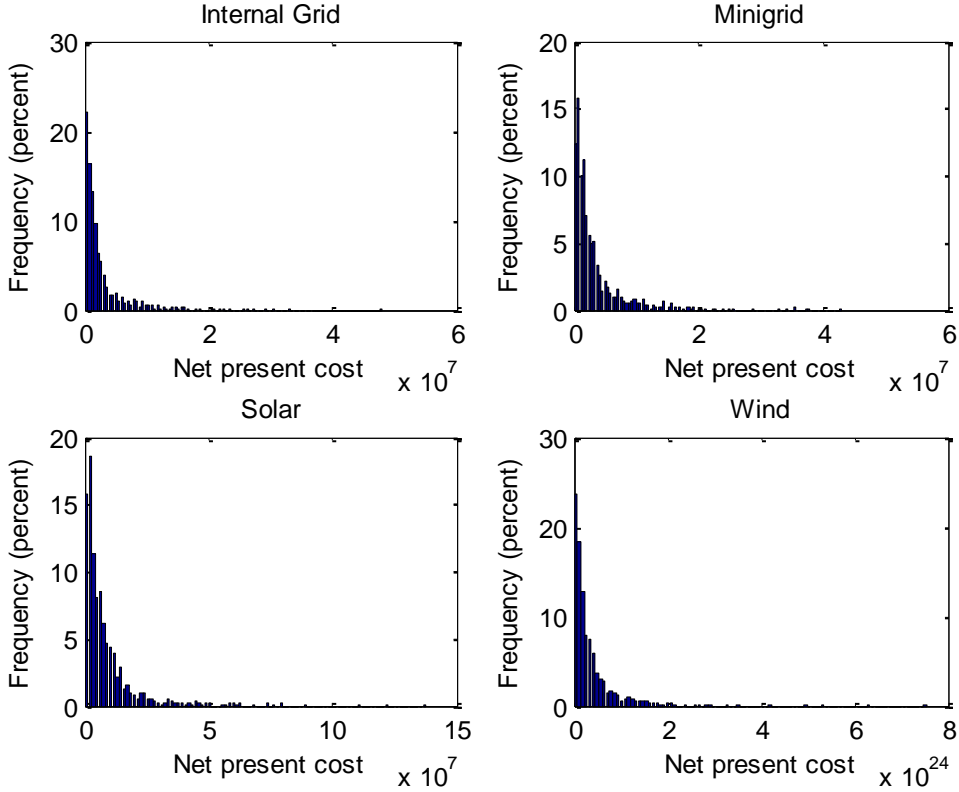


Figure 5: Distribution of net present costs of electrification technologies in Ghana

Mean and median costs of the distributions are \$2.93million and \$1.43million for internal grid costs, \$3.44million and \$1.81million for mini-grid costs, \$8.86million and \$5.04million for solar technology costs and prohibitively high for wind technology costs. The low mean and median of the grid cost distribution is because internal grid costs only are considered; capital and recurrent costs of primary MV connections needed to network nodes into the existing grid are not factored. The prohibitively high mean and median of the wind cost distribution is as a result of very low wind resource for settlements in Ghana (i.e. less than 2m/s for over 90% of settlements in the data). For wind technology to be viable, substantial wind resource is required ($> 6\text{m/s}$). NREL (2011) estimates that only 0.2% of Ghana’s total land mass has this level of wind potential.

3.2 Evaluation measures

The example in Section 0 above showed that the PA algorithm may generate solutions which differ in terms of cost and also the extent of the use of off-grid technologies. To capture both dimensions we judge the performance of the PA algorithm relative to the MINLP model using the mean absolute error and mean absolute percentage error in terms of the ‘total cost’ and the ‘number of nodes connected to the grid’. More formally, let t represent the t^{th} Monte Carlo trial and PA_t and $MINLP_t$ represent a measured solution metric for the respective methods, then we define

$$MAE = \frac{1}{n} \sum_t |PA_t - MINLP_t| \quad (11)$$

$$MAPE = \frac{1}{n} \sum_t \frac{|PA_t - MINLP_t|}{MINLP_t} \times 100 \quad (12)$$

where n is the number of simulations and MAE and $MAPE$ are the absolute error and the percentage absolute error respectively (Swanson et al., 2011)¹⁰.

Spatial factors influence electrification costs (Zvoleff et al., 2009) and may influence the relative cost effectiveness of both methods. We therefore test their sensitivities to two spatial factors known to be typical of unelectrified nodes in developing countries namely 1) the degree of spatial dispersion between the simulated nodes and 2) their degree of remoteness from the existing grid. Greater dispersion and remoteness increase electrification costs by increasing the length of the MV lines needed to connect nodes. Let $distance_{grid}(i)$ represent distance from an unelectrified node i to the grid. For any simulated scenario, we calculate degree of remoteness of nodes from the existing grid and degree of dispersion as follows.

$$remoteness_from_grid = \sum_i distance_{grid}(i) \quad (13)$$

$$dispersion = \sum_{i,j} distance(i,j) \quad (14)$$

where $distance(i,j)$ is Euclidean distance between nodes as previously defined.

4 Results and discussion

4.1 Cost performance

Table 3 reports the relative overall cost results and also results disaggregated by the two measures used to capture the impact of spatial effects, namely, the spatial dispersion and remoteness metrics. Columns 1 to 3 show the frequency of times the MINLP model outperforms the heuristic PA algorithm. Overall in 87.6% of the trials, the MINLP total cost is lower. However, column 4 and 5 results show that although the MINLP model yields lower costs in a majority of our simulation scenarios, the overall cost difference between both methods is small. The mean absolute error is \$365,000, which constitutes a mean absolute

¹⁰ By definition $PA_t - MINLP_t \geq 0$ for the ‘total cost’ solution metric hence the use of the absolute measures is strictly unnecessary and a simple average would suffice. Similarly the results would show that $PA_t - MINLP_t \leq 0$ in terms of the ‘number of nodes connected to grid’ solution metric in all cases hence use of absolute measures is again strictly unnecessary.

percentage error of 0.7% only. However, as shown in Figure 6, the degree to which the PA algorithm is outperformed by the MINLP model in certain scenarios can be significant, with the maximum percent absolute error for cost found to be 3.7% (representing around \$1.8 million).

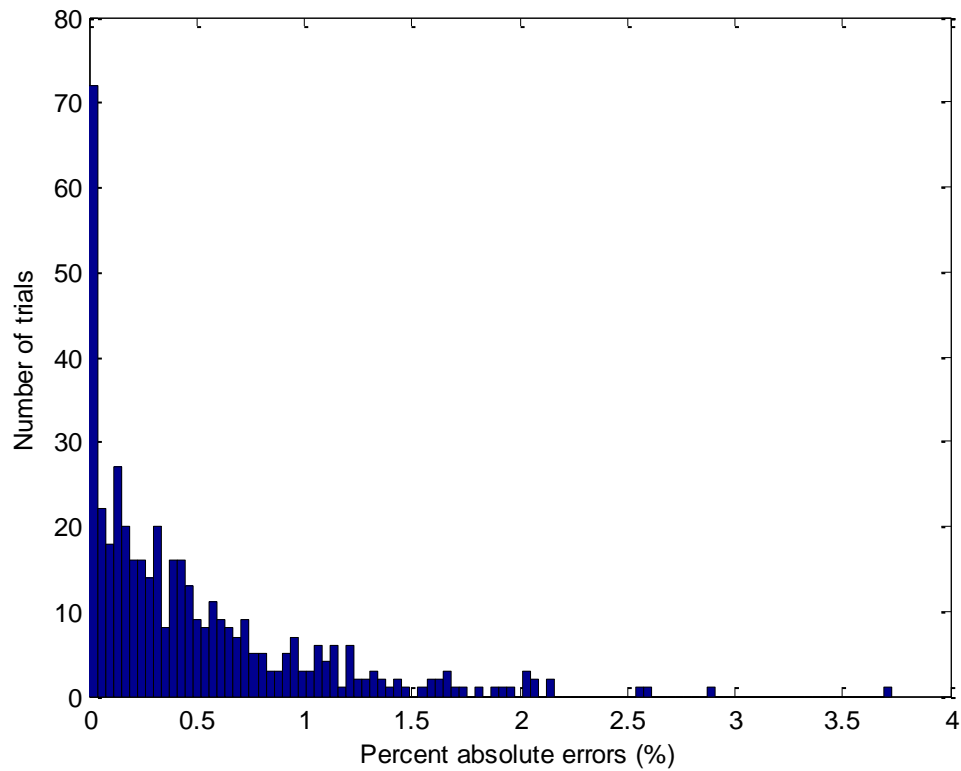


Figure 6: Full range of percent absolute errors in cost across all trials

	Freq. of lower cost solutions in MINLP	Freq. of lower cost solutions in PA	Freq. of equal costs in both	MAE, \$'000	MAPE	Maximum AE_t , \$ million	Maximum PAE_t	Std. dev. MINLP, \$ million	Std. dev. PA, \$ million
Overall simulation results	87.6%	0.0%	12.4%	365	0.7%	1.80	3.7%	14.53	14.53
Distribution of results by spatial dispersion metric									
Up to 25th percentile	78.5%	0.0%	21.5%	146	0.3%	0.75	2.0%	13.81	13.80
25 - 50th percentile	87.0%	0.0%	13.0%	311	0.6%	0.95	2.0%	14.80	14.79
50 - 75th percentile	91.7%	0.0%	8.3%	401	0.8%	1.36	2.6%	14.09	14.08
> 75th percentile	93.5%	0.0%	6.5%	505	1.0%	1.80	3.7%	15.00	15.03
Distribution of results by remoteness from grid metric									
Up to 25th percentile	78.5%	0.0%	21.5%	172	0.4%	0.75	2.0%	13.50	15.00
25 - 50th percentile	88.0%	0.0%	12.0%	249	0.5%	0.95	2.0%	14.67	14.66
50 - 75th percentile	88.0%	0.0%	12.0%	351	0.7%	1.08	2.6%	15.11	15.06
> 75th percentile	96.3%	0.0%	3.7%	560	1.1%	1.80	3.7%	14.63	14.68

Table 3: Simulation results comparing relative costs of the MINLP model and the PA algorithm, and the sensitivities of their relative costs to spatial factors.

While the overall extent of errors in terms of cost between both methods is small, the disaggregation of the results in Table 3 by spatial dispersion and remoteness from the grid do suggest that this error does increase with these two factors. As spatial dispersion of the unelectrified nodes increases in the simulations, MAPE monotonically increases ranging from average of 0.3% for below the 25th percentile of the dispersion metric to 1.0% for above the 75th percentile. Similarly, as remoteness from the grid increases, MAPE monotonically increases from 0.4% for below the 25th percentile of the remoteness metric to 1.1% for above the 75th percentile. These monotonic increases are also true of other measures such as MAE, maximum AE_i and maximum PAE_i . Figure 7 shows the full range of the distribution of percent absolute error, disaggregated by the two spatial factors. Correlation coefficient between percent absolute errors with dispersion and remoteness are +0.38 and +0.44 respectively. The lines of best fit are also shown to emphasise these associations.

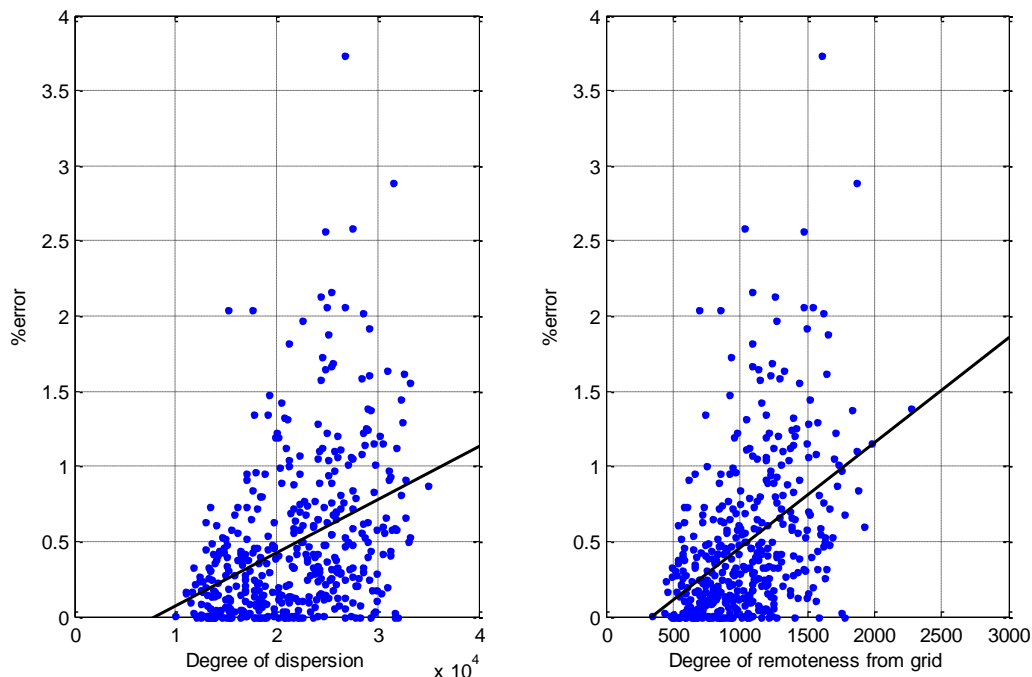


Figure 7: Distribution of percent absolute errors in cost, separated by spatial metrics

Given that unelectrified nodes in developing countries are typically dispersed and remote from the existing grid, the deteriorating relative performance of the PA algorithm is potentially an important consideration in electricity planning.

4.2 Grid versus off-grid performance

Although cost minimisation is the primary objective of both methods, their relative performance in terms of their propensity to allocate grid or off-grid sources to unelectrified nodes is a potentially important planning issue. Bhattacharyya (2013) notes that developing countries have traditionally only viewed off-grid sources as inferior temporary solutions prior to grid electrification whilst Urpelainen (2014) recommends that ‘grid expansion should be the ultimate goal in national electricity planning’. Evidence also suggests that unelectrified communities in developing countries prefer grid electrification and exert political influence to

facilitate this preference (Bawakyillenuo, 2007, Brown and Mobarak, 2009, Min, 2011). Greater allocation of nodes to the grid is therefore seen as desirable as it leads to more solutions which are often more politically acceptable (Abdul-Salam and Phimister, 2016).

Table 4 shows the simulation results relative to the allocation of grid nodes by both algorithms. Columns 1 to 3 show that on no occasion did the PA algorithm award more grid nodes than the MINLP model. To the contrary, the MINLP model awards more grid nodes in 74.3% of the trials. The MAE is therefore interpreted as the average number of grid nodes the MINLP model allocates more than the PA algorithm; MAPE is interpreted as the percentage level of this number. On average, there are 25.2% (MAPE) more nodes allocated to the grid in the MINLP model than in the PA algorithm.

To explore this property further, we distribute results by the two spatial factors that affect electrification costs. Distributing the results by the degree of spatial dispersion metric shows that increasing dispersion leads to a deteriorating performance of the PA algorithm relative to the MINLP model. MAPE in the trials with the lowest degree of dispersion (i.e. up to the 25th percentile of degree of dispersion metric) is 14.1%. This rises to 32.8% in the trials with the highest degree of dispersion (i.e. above 75th percentile). Similarly, distributing the results by the metric of nodal remoteness from the grid shows that MAPE increases from 14.9% in the trials with the closest affinity of nodes to the grid (i.e. up to the 25th percentile of degree of remoteness metric) to 32.1% in the trials with nodes further placed from the existing grid (i.e. above 75th percentile). These monotonic increases are also largely true of other measures such as MAE , maximum AE_i and maximum PAE_i .

	Freq. of more grid nodes in MINLP	Freq. of more grid nodes in PA	Freq. of equal no. of grid nodes	MAE	MAPE	Maximum AE_t	Maximum PAE_t	Std dev (MINLP)	Std dev (PA)
Overall simulation results	74.3%	0.0%	25.7%	2.74	25.2%	10.00	83.3%	2.82	3.32
Distribution of results by spatial dispersion metric									
Up to 25th percentile	58.9%	0.0%	41.1%	1.85	14.1%	7.00	45.5%	2.17	2.32
25 - 50th percentile	71.3%	0.0%	28.7%	2.89	24.6%	10.00	81.8%	3.01	3.42
50 - 75th percentile	80.2%	0.0%	19.8%	2.85	26.0%	8.00	77.8%	2.39	2.79
> 75th percentile	86.8%	0.0%	13.2%	3.22	32.8%	10.00	83.3%	2.94	3.27
Distribution of results by remoteness from grid metric									
Up to 25th percentile	57.9%	0.0%	42.1%	1.79	14.9%	6.00	71.4%	2.50	2.63
25 - 50th percentile	75.0%	0.0%	25.0%	2.62	22.0%	10.00	66.7%	2.75	3.07
50 - 75th percentile	75.7%	0.0%	24.3%	3.11	28.5%	10.00	81.8%	3.11	3.50
> 75th percentile	88.6%	0.0%	11.4%	3.22	32.1%	9.00	83.3%	2.54	3.12

Table 4: Simulation results comparing relative propensities of the MINLP model and the PA algorithm to assign nodes to the grid rather than off-grid, and the sensitivity of this propensity to spatial factors.

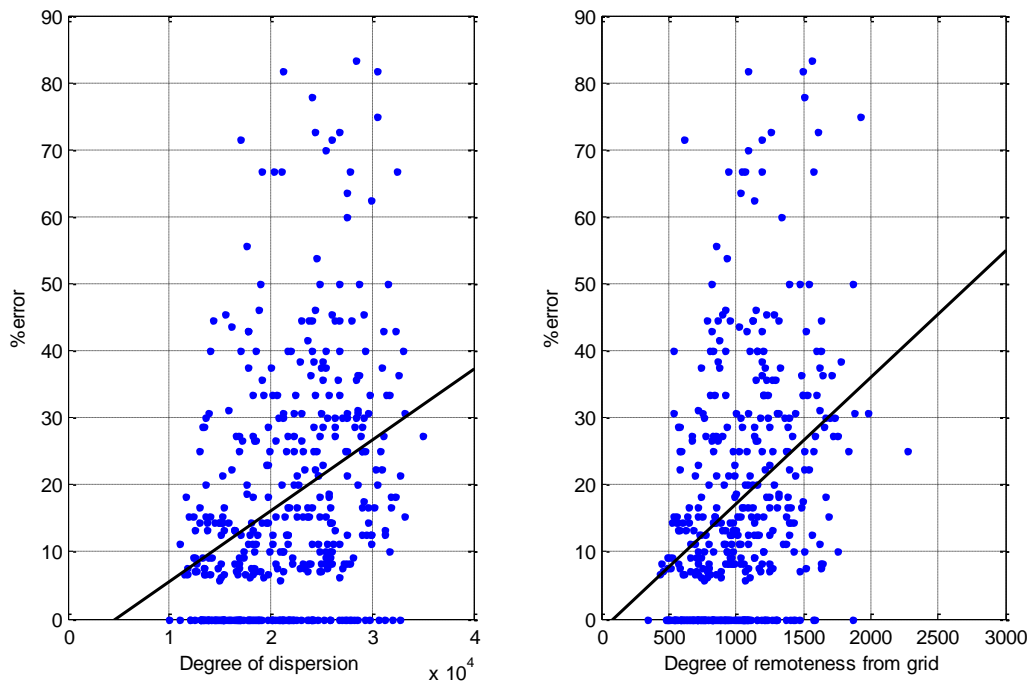


Figure 8: Distribution of percent absolute errors in number of grid assigned nodes, separated by spatial metrics.

Figure 8 shows the full range of the distribution of percent absolute error, disaggregated by the two spatial factors. Correlation coefficient between percent absolute errors with dispersion and remoteness are $+0.34$ and $+0.36$ respectively. The lines of best fit are again shown to emphasise these associations.

4.3 Sub-regional planning example

Finally, we implement the MINLP model to a defined case study of Ghana in order to demonstrate how the model might be used for small sub-regional planning problems. We use the Ghana data described in Section 0, as adapted from Abdul-Salam (2015). Figure 9 shows the MINLP model solution for a 40 node problem in the North-West of Ghana as shown by the spatial indicators.

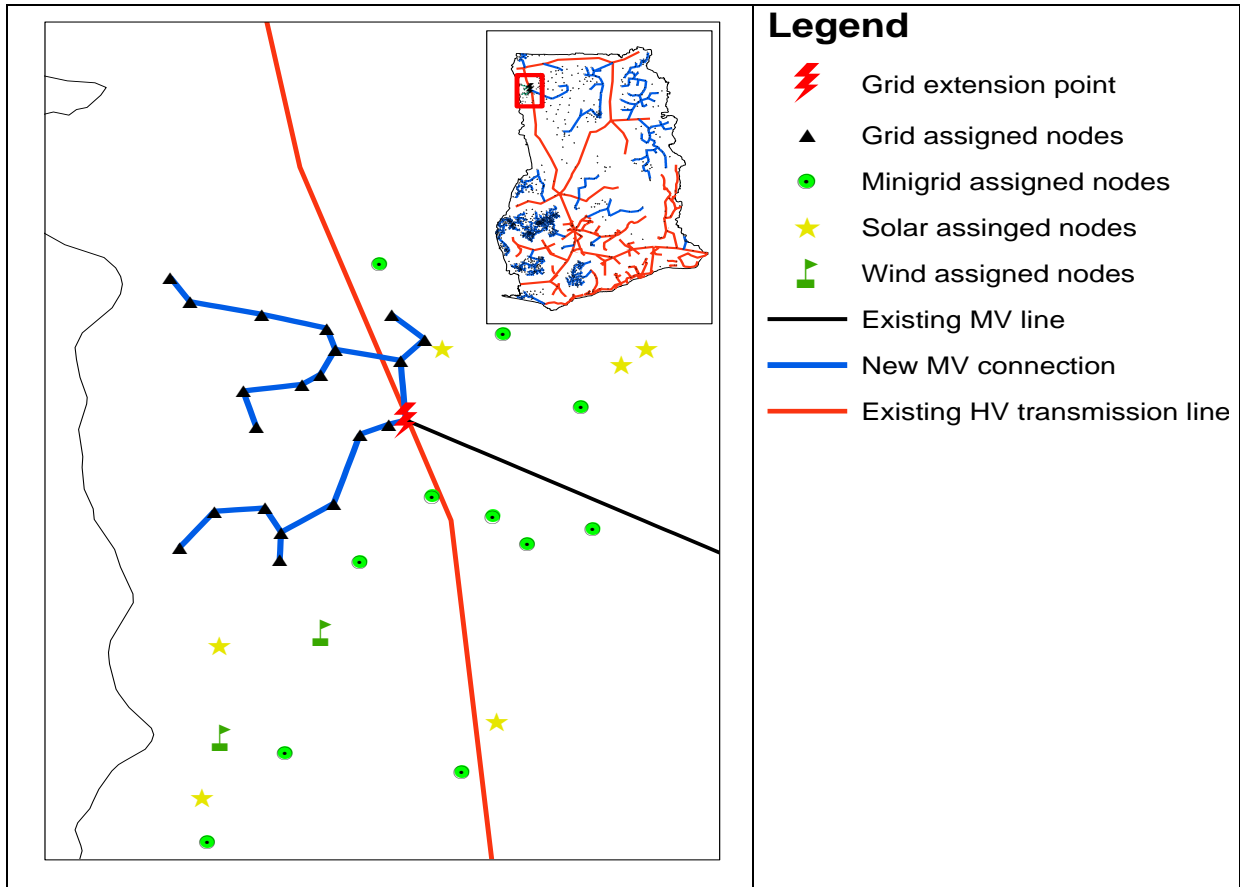


Figure 9: Application of the MINLP model to a small sub-regional problem involving 40 nodes.

In the above solution, the MINLP model connects 21 nodes to the central grid and assigns 6, 2 and 11 nodes for standalone solar, wind and autonomous mini-grid systems respectively for a total electrification cost of \$23.19 million over 10 years. The PA algorithm also yielded the exact same solution. The equality in solutions supports our overall simulation findings, that the PA algorithm generally yields cost effective solutions that are close to the global optimum in the MINLP model. Whilst there was no cost advantage to use of the MINLP model in this instance, we maintain that use of it in sub-regional planning is preferred. As our simulations showed, its use might result significant cost savings in some situations. It also has greater likelihood of connecting nodes to the grid, leading to greater political acceptance of the solutions it yields.

5 Conclusion

In the context of developing countries, electricity planning involving grid and off-grid sources has been promoted as the best means of achieving cost effective universal access to electricity in the short to medium term. However, large scale universal electricity planning involving these sources with consideration for the economic and networking aspects of electrification is complex. As such heuristic algorithms have traditionally been used as computationally feasible means of yielding cost effective planning solutions. With these

algorithms however, global optimality is not guaranteed. We develop an MINLP model of the underlying problem. By definition, the model yields optimum solutions and provides a basis for testing the relative performance of the widely used heuristic algorithm due to Parshall et al. (2009) (referred to here as the PA algorithm). Due to the computational intensity of the MINLP model however, testing is done with a Monte Carlo simulation procedure involving small scale trials of a simplified version of the underlying planning problem.

Overall the results show that in general, the typical difference in the cost effectiveness of both methods is relatively small. However, examining the effect of spatial factors that are typical of unelectrified settlements in developing countries, we do find that the relative performance of the PA algorithm deteriorates with these factors. For example, in one scenario, the expected saving from use of the MINLP model rather than the PA algorithm is about \$1.8 million. This level of saving would be sufficient to fund electricity access for multiple additional small communities in our data. Additionally, the PA algorithm solution in some situations overstates the role of off-grid electrification sources which may imply solutions that are potentially more expensive and/or less politically acceptable for the communities involved.

Given our results, the MINLP model may be used for sub-regional versions of the universal electricity planning problem involving few nodes. However, given that the PA algorithm typically yields solutions that are close to the global minimum, our results imply that it can be used with a degree of confidence in most situations. Further it would seem to provide a reasonable benchmark to validate other heuristic algorithms for large scale electricity planning (see Abdul-Salam and Phimister, 2016). However given that its relative performance deteriorates in dispersed and remote regions, some care should be taken when used in such situations. Finally, future work developing methods allowing the solution of the MINLP model in large scale problems would appear to be of value, not least as it would permit the exploration as to whether the results obtained here are sustained at that scale.

Acknowledgement

The first author would like to acknowledge the University of Aberdeen and the Henderson Economics Research Fund for funding his PhD studies in the period 2011-2014 which formed the basis for the research presented in this paper. The first author would also like to acknowledge the Macaulay Development Trust which funds his postdoctoral fellowship with The James Hutton Institute, Aberdeen, Scotland. The authors thank two anonymous referees for valuable comments and suggestions on earlier versions of this paper. All usual caveats apply.

References

- Abdul-Salam, Y., 2015. "Access to electricity in Sub Saharan Africa: Modelling the importance and adoption of off-grid renewables." (Doctoral dissertation, University of Aberdeen).
- Abdul-Salam, Y, and Phimister, E., 2016 "The politico-economics of electricity planning in developing countries: A case study of Ghana." *Energy Policy* 88: pp. 299-309.
- Afripop, 2012. "Population GIS data for Ghana." Available at <http://www.afripop.org/> (accessed January, 2012).
- Akpan, U., 2015. "Technology Options for Increasing Electricity Access in Areas with Low Electricity Access Rate in Nigeria." *Socio-Economic Planning Sciences*, Volume 51, Pages 1-12, ISSN 0038-0121. Available at <http://dx.doi.org/10.1016/j.seps.2015.05.001>
- Amador, J., Domínguez, J., 2005. "Application of geographical information systems to rural electrification with renewable energy sources". *Renewable Energy*, 30(12), 1897-1912.
- Bawakyillenuo, S. 2007. "Rural electrification in Ghana: Issues of photovoltaic energy technology utilization." (Doctoral dissertation, University of Hull).
- Bazilian, M., Welsch, M., Divan, D., Elzinga, D., Strbac, G., Howells, M., Jones, L., Keane, A., Gielen, D., Balijepalli, V.S.K.M. and Brew-Hammond, A., 2011. "Smart and just grids: opportunities for sub-Saharan Africa." London: The Energy Futures Lab, Imperial College.
- Bhattacharyya, S. C., 2013. "To regulate or not to regulate off-grid electricity access in developing countries." *Energy Policy*, 63, 494-503.
- Bhattacharyya, S.C., 2006. "Energy access problem of the poor in India: is rural electrification a remedy?" *Energy Policy* 34 (18), 3387–3397.
- Brown, D. S., Mobarak, A. M., 2009. "The transforming power of democracy: regime type and the distribution of electricity." *American Political Science Review*, 103(02), 193-213.
- Buys, P., Deichmann, U., Meisner, C. M., That, T. T., Wheeler, D., 2007. "Country stakes in climate change negotiations: two dimensions of vulnerability." *World Bank Policy Research Working Paper*, (4300).
- Deichmann, U., Meisner, C., Murray, S., Wheeler, D., 2011. "The economics of renewable energy expansion in rural sub-saharan Africa." *Energy Policy*, 39(1), 215-227.
- Demierre, J., Bazilian, M., Carbajal, J., Sherpa, S., & Modi, V., 2015. "Potential for regional use of East Africa's natural gas." *Applied Energy*, 143, 414-436.
- Ferrer-Martí, L., Pastor, R., Capó, G. M., & Velo, E., 2011. "Optimizing microwind rural electrification projects. A case study in Peru." *Journal of Global Optimization*, 50(1), 127-143.
- Goldemberg, J., 2000. "World Energy Assessment: Energy and the Challenge of Sustainability." United Nations Publications, New York, NY.

- International Energy Agency (IEA), 2013. "World Energy Outlook (WEO)." Available: <http://www.worldenergyoutlook.org/publications/weo-2013/> (accessed February 2015)
- Kemausuor, F., Adkins, E., Adu-Poku, I., Brew-Hammond, A., Modi, V., 2014. "Electrification planning using network planner tool: The case of Ghana." *Energy for Sustainable Development*, 19, 92-101.
- Khandker, S., Douglas F., and Hussain, A., 2009. "Welfare impacts of rural electrification: a case study from Bangladesh." *World Bank Policy Research Working Paper Series*.
- Kruskal, J. B., 1956. "On the shortest spanning subtree of a graph and the traveling salesman problem." *Proceedings of the American Mathematical Society.*, 7(1), 48-50.
- Lambert, T. W., Hittle, D. C., 2000. "Optimization of autonomous village electrification systems by simulated annealing." *Solar Energy*, 68(1), 121-132.
- Levin, T., Thomas, V. M., 2012. "Least-cost network evaluation of centralized and decentralized contributions to global electrification." *Energy Policy*, 41, 286-302.
- Maristes, H., 2011. "The Rural Electrification Senegal (ERSEN) project: electricity for over 90,000 persons." PERACOD, Dakar, Senegal. Accessed at: <http://www.peracod.sn/IMG/pdf/ENRbases-technologies.pdf> (Feb 2016)
- Min, B., 2011. "Electrifying the Poor: Distributing Power in India." *Ann Arbor 1001*: 48109-1045.
- National Renewable Energy Laboratory (NREL), 2011. "Ghana Wind Energy Resource Mapping Activity." Available at <http://en.openei.org/doe-opendata/dataset> (accessed August 2012)
- Nguyen, K.Q., 2007. "Alternatives to grid extension for rural electrification: Decentralized renewable energy technologies in Vietnam." *Energy Policy*, **35**(4), pp. 2579-2589.
- Ohiare, S., 2015. "Expanding electricity access to all in Nigeria: a spatial planning and cost analysis." *Energy, Sustainability and Society*, 5(1), 1-18.
- Painuly, J. P., Fenhann, J. V. , 2002. "Implementation of renewable energy technologies-opportunities and barriers. Summary of country studies." Risø National Laboratory. UNEP Collaborating Centre on Energy and Environment.
- Parshall, L., Pillai, D., Mohan, S., Sanoh, A., Modi, V., 2009. "National electricity planning in settings with low pre-existing grid coverage: Development of a spatial algorithm and case study of Kenya." *Energy Policy*, 37(6), 2395-2410.
- Prim, R.C., 1957. "Shortest connection networks and some generalizations." *Bell system technical journal*, **36**(6), pp. 1389-1401.
- Ranaboldo, M., Ferrer-Martí, L., García-Villoria, A., & Moreno, R. P., 2013. "Heuristic indicators for the design of community off-grid electrification systems based on multiple renewable energies." *Energy*, 50, 501-512.

Sanoh, A., Parshall, L., Sarr, O. F., Kum, S., Modi, V., 2012. "Local and national electricity planning in senegal: Scenarios and policies." *Energy for Sustainable Development*, 16(1), 13-25.

Sempértegui, R., Bautista, J., Grino, R., & Pereira, J.. 2002. "Models and procedures for electric energy distribution planning. A review." IFAC 2002.

Solar and Wind Resource Assessment (SWERA), 2011. Available at <http://en.openei.org/wiki/SWERA/Data> (accessed December, 2011)

Swanson, D. A., Tayman, J., & Bryan, T. M., 2011. "MAPE-R: a rescaled measure of accuracy for cross-sectional subnational population forecasts." *Journal of Population Research*, 28(2-3), 225-243.

Szabó, S., Bódis, K., Huld, T., & Moner-Girona, M., 2013. "Sustainable energy planning: Leapfrogging the energy poverty gap in Africa." *Renewable and Sustainable Energy Reviews*, 28, 500-509.

Urpelainen, J., 2014. "Grid and off-grid electrification: an integrated algorithm with applications to India." *Energy for Sustainable Development*, 19, 66-71.

World Bank, 2007. "Providing Electricity to Poor Rural Provinces of Lao PDR." Available at <http://documents.worldbank.org/curated/en/2007/04/24014041/providing-electricity-poor-rural-provinces-lao-pdr> (accessed February 2016)

World Bank, 2008. "The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits." World Bank. Available at https://ieg.worldbankgroup.org/Data/reports/rural_elec_full_eval.pdf (accessed February 2016)

World Bank, 2013. "Implementation and Completion Report: Republic of Peru Rural Electrification Program." Available at http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2014/01/27/000442464_20140127095748/Rendered/PDF/ICR23580P090110C0disclosed010230140.pdf (accessed February 2016)

World Bank, 2015. "Access to Electricity (% of population)." Available at <http://data.worldbank.org/indicator/EG.ELC.ACCS.ZS> (accessed May 2015)

Zvoleff, A., Kocaman, A. S., Huh, W. T., & Modi, V., 2009. "The impact of geography on energy infrastructure costs." *Energy Policy*, 37(10), 4066-4078.