A regional approach to optimizing the location of rural handpumps

Omar S. Hopkins

ABSTRACT

Many rural water supply projects in sub-Saharan Africa are based on the installation of public handpumps. One of the key benefits of these projects is distance and time savings. Surprisingly, references to rural water planning rarely provide systematic approaches to optimizing distance-related benefits. This paper develops a conceptual model to identify the number and location of point sources that maximizes benefits to consumers, thereby serving as an aid to decision makers in identifying good alternatives. The proposed model is based on willingness to pay, a location model to identify optimal locations of sources, and a cost–benefit analysis. The model shows that as the number of sources increases, the distance between households and sources decreases but the user fee must increase to generate the revenue required to maintain them. Higher fees will dissuade households from using the point sources and hence reduce the aggregate distance savings that accrue. This suggests that there is an optimal number and location of point sources.

Key words | facility location models, optimization, planning, rural water

Omar S. Hopkins
Director,
Millennium Challenge Corporation,
875 15th Street NW,
Washington,
DC 20005,
USA
E-mail: *hopkinsos@mcc.gov*

INTRODUCTION

Over a billion people lack adequate access to potable water supply which has adverse effects on human health, productivity, and education of children; but perhaps the most onerous is the time and effort expended collecting water from distant sources (Hutton & Haller 2004). Sustainability is a key challenge in rural water projects as evidenced by the high failure rates (RWSN Executive Steering Committee 2010). A key element of sustainability is financial performance which is achieved when consumers receive the level of service for which they are willing to pay.

In many rural areas, where households are poor and the costs of constructing a distribution network to reach a dispersed population are high, planners may choose to implement a network of rural handpumps which triggers decisions on how many to construct and where to locate them. The current state of practice in many countries calls for locating public access points within some maximum distance from user households and to limit the number of persons relying on a single source to ensure that some minimum level of service is achieved. Typically point sources are located to ensure that the furthest household is no more than 250–500 m and that no more than 250–1,000 persons use any one point source (Carl Bro International 1997; Ministry of Water & Livestock Development 1997) – which essentially acknowledges that households incur a cost if the source is far away or if they must wait for a long time at the source to collect water. These guidelines, which identify minimum service standards, do not address the issue of how to identify the optimal number and location of sources nor how to maximize the net benefits to a community. Nonetheless, community planners must not only identify the best technology or technologies (handpumps, stand posts, etc.) for a community but also determine the 'number … and precise location of the facilities' (van Wijk-Sijbesma & Smet 2002). What literature exists on siting point sources is often focused on understanding hydrogeological constraints (Carter *et al.* 2014). There exists a significant gap in the literature relating to the design of rural water supply systems to optimize the distance-related time savings.

doi: 10.2166/washdev.2015.128

Over the past decade or more, there have been increasing efforts to use geospatial information and tools to help inform planning and to support monitoring and evaluation in the rural water sector. However, substantial challenges to implementing geospatial strategies exist: access to reliable data, ability to share that data, and using that data to facilitate planning (Pearce & Howman 2012; Welle 2010).

Recent efforts demonstrate the challenges and opportunities in improving data collection (Giné-Garriga *et al.* 2013) and analysis (Jiménez & Pérez-Foguet 2011). Some simple open source tools exist to facilitate analysis based on coverage (WaterAid 2010), but they do not generally include robust tools to analyze optimal service strategies. Some GIS packages provide tools to help planners locate public facilities, including potentially handpumps (ESRI 2012). However, one of the major limitations of these tools is that they are focused on optimal location and not optimal number of handpumps in a given region. Many funding agencies, including the Millennium Challenge Corporation (MCC), make investment decisions based on a range of factors, including economic rates of return (MCC 2013a). These agencies might be interested in identifying good, if not near optimal, levels of investment. The proposed model provides a strategy for making that estimate and additionally provides information about the location of handpumps based on three components: willingness to pay estimates, a facility location model, and cost–benefit analysis (CBA).

METHODS AND DATA

The proposed model was tested in a portion of the Mozambican District of Mogovolas in the province of Nampula; the residential spatial distribution was taken from Google Earth images (Burke 2014). This area is within the zone where MCC previously funded a rural water project (Governments of the United States & Mozambique 2007). A square with 10 km sides was defined and 50 housing clusters, or neighborhoods, were identified, located, and their populations estimated. The centroid of the illustrative regional area is roughly located at longitude 39.4 °E and latitude 15.8 °S. Based on the satellite images, the author estimates that this 100 km^2 area has a population of 3,500 persons (or 835 households assuming 4.2 persons per household) which is roughly equivalent to the population density of the province as a whole less the principal population centers of Nacala Porto and Cidade de Nampula (Instituto Nacional de Estatistica 2014). The households were divided into 50 neighborhoods. Figure 1 illustrates the locations of the clusters, the underlying population density heat map, and a segment of the Nampula–Angoche road.

Organization of the population data for each neighborhood and distances between neighborhood centroids is shown in Table 1. The coordinates were used to estimate simple straight line distances between communities; straight line distances are a reasonable proxy for actual distances (Ho *et al.* 2014).

In Mozambique, cost was the second most significant reason given for not using an improved source, following distance (Hall *et al.* 2014). The willingness-to-pay (WTP) model enables predictions about the fraction of households willing to pay for access to an improved rural water supply system. WTP can be estimated using a variety of methods including, but not limited to, contingent valuation and revealed preference methods (Nauges & Whittington 2009). Unfortunately, in this project no willingness to pay estimate was prepared, so for the purpose of the model it

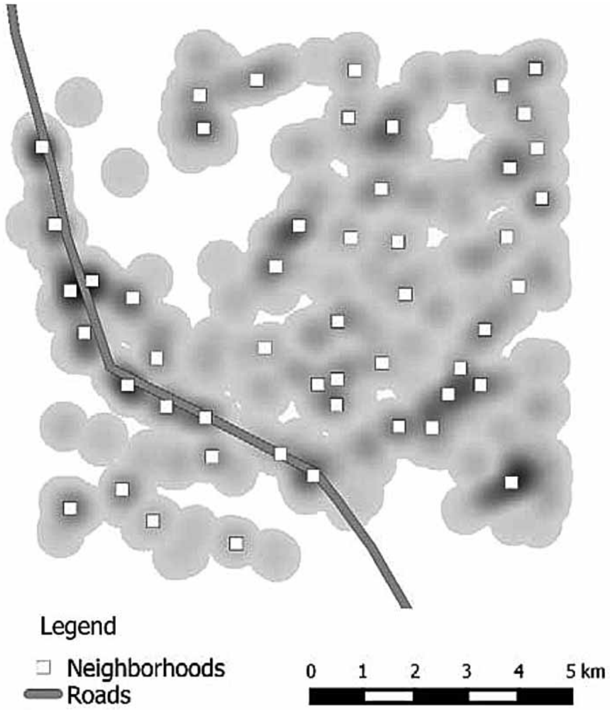

Figure 1 | Map of neighborhood locations.

Table 1 | Distance and population matrix

	Neighborhood ID	**Long. (°E)**	**Lat. (°S)**	**Pop (HH)**	**From (row) – to (col) distance in km**					
From Neighborhood …					1	2	3	…	49	50
	1	39.3693	15.7734	132	0	1.5	2.8	…	6.8	6.4
	2	39.3716	15.7868	108	1.5	0	1.3	…	5.3	5.9
	3	39.3743	15.7983	156	2.8	1.3	0	…	4.0	5.7
	…	…	…	…	…	…	…	…	…	…
	49	39.3836	15.8328	52	6.8	5.3	4.0	…	0	6.7
	50	39.4247	15.7891	24	6.4	5.9	5.7	…	6.7	0

is assumed that if the typical monthly charge was 30 Meticais (roughly US$1) per month then about 30% of households are not willing to pay to use the new point sources and that if the price was zero there would be no financial barrier to use. Equation (1) illustrates the linear relationship between the fee (Fee) in Meticais and the fraction of households willing to pay (WTP) that fee.

$$\mathrm{WTP} = 1.00 - 0.01\ \mathrm{Fee} \tag{1}$$

One of the key elements of sustainability is the ability of consumers to generate revenue to cover the cost of an improved water supply. Planners are therefore interested in knowing if a scheme is financially feasible, which requires estimating the ability of a community to generate revenue. Equation (2) illustrates that the expected revenue per household (R) that can be generated is simply the product of the fraction of households willing to pay fee and that fee.

$$R = \mathrm{WTP} \times \mathrm{Fee} \tag{2}$$

For example, from Equation (1) when the user fee is set to 30 Meticais/month/household, approximately 70% of households are predicted to pay that fee – and 30% will not and will continue to use their traditional sources. The expected revenue per household (Equation (2)) is then predicted to be 21 Meticais per month per household; total revenue is the product of expected revenue and the number of households. Within the region of analysis there is a maximum amount of revenue that can be generated and hence a maximum number of improved point sources that is affordable (Wedgwood & Sansom 2002). The expected revenue, $E[R]$, curve (Figure 2) has a clear maximum when approximately 50% of households are WTP the user fee of 50 Meticais/month. If WTP is either above or below the level of 50%, revenues decline. In general, planners should only consider the region to the left of the maximum revenue because for any given level of revenue it is better to serve more households rather than less.

Assume that the costs, typically limited to the operation and maintenance (O&M), faced by users are approximately

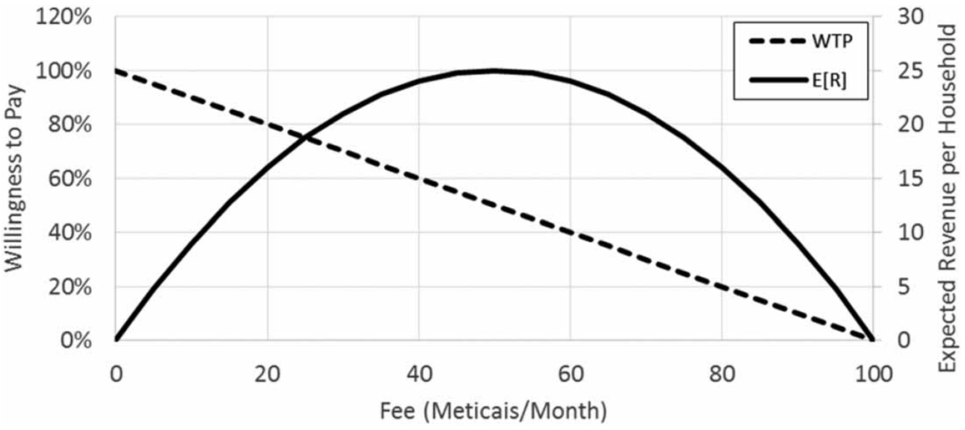

Figure 2 | WTP, user fees, and revenue.

$500/source per year; this could cover actual parts and labor for repairs, the cost of an attendant, and honorariums for the water committee (WashCost 2012). If the fees are set so as to generate enough revenue to exactly cover the O&M costs, an analyst can calculate the fee required to support any number of improved sources. The maximum number of point sources that are financially feasible is a function of the maximum expected revenue per household, which occurs when the user fee is set to 50 Meticais (US$1.65) per month, and half the households are willing to pay that fee to use an improved source. Assuming an operating cost of $500 per year, this community of 3,500 individuals or 835 households could support up to 16 handpumps; therefore the alternatives are to construct between 1 and 16 new handpumps and the planning question is how many to build and where to locate them.

Location models are a type of optimization model used to solve problems that seeks to match points of demand with points of supply. The p-median location model (ReVelle & Swain 1970) is used to locate schools, police stations, and health clinics to minimize the population-weighted distance between points of demand and points of supply. Hopkins *et al.* (2004) illustrate how this model can be used to locate rural handpumps to minimize the distance between consumers and point sources. Location models are based on the assumptions that: (1) candidate locations are drawn from a homogeneous demand area; (2) demand, although distributed within an area, can be represented at a single point; and (3) beneficiaries will use their assigned facility. The p-median model has several advantages: (1) it identifies the Pareto optimal solution such that no neighborhood can reduce its travel distance for collecting water from improved sources without making another neighborhood worse off; (2) it has an optimal solution where the improved sources are situated at the candidate locations; (3) it is easily solved by linear programming; and (4) it requires data that are readily available from published sources. The objective (Equation (3)) of the p-median model is to minimize the weighted distance between households and improved sources, where pi is the number of households in neighborhood i, dij is the one-way distance that a household in neighborhood i travels to collect water from a source in neighborhood j, and $xij = 1$ if households in neighborhood i are assigned to use an improved source in neighborhood j (0 otherwise). Equation (4) requires that households in each neighborhood must be assigned to an improved source. If $xjj = 1$, then households in neighborhood j collect water from their own neighborhood. Hence, Equation (5) restricts the total number of improved sources to P. Finally, if some neighborhood i is assigned by the computer to collect water from neighborhood j, then households in j must self-assign to j, which is accomplished by Equation (6).

$$\text{Min} \sum_{i=1}^{N} \sum_{j=1}^{N} p_i d_{ij} x_{ij} \tag{3}$$

$$\sum_{j=1}^{N} x_{ij} = 1, \ \forall i \tag{4}$$

$$\sum_{j=1}^{N} x_{jj} = P \tag{5}$$

$$x_{jj} > x_{ij}, \ \forall i \forall j \tag{6}$$

US Government agencies, like the MCC, often use CBA or variations to inform decision making. The constituent elements of CBA are costs and benefits streams, when they accrue in time, and the social discount rate. Costs and benefits are discounted based on how far in the future they are expected to occur and when summed are the net present value (NPV) of the project. The alternative with the highest NPV is preferred and corresponding locations as determined by the p-median model are optimal.

As a matter of policy MCC uses a 10% discount rate and a 20-year planning horizon unless a longer or shorter planning period is deemed appropriate (MCC 2013b). The benefits are derived from the results of the p-median model which is evaluated with the number of improved sources ranging between zero and the maximum number that is financially feasible.

The use of CBA requires that the benefit streams can be expressed in monetary terms. In the case of rural water projects, it is assumed that distance benefits are the primary driver for changes in NPV of the alternative. Every household consumes water, albeit perhaps from different sources, and all households have costs associated with collecting water. As distances are a principal driver of benefits, one of the challenges is to estimate the value of distance savings. The key assumptions in this paper are: an average walking speed of 3 km per hour, an 8-hour work day, a value of time equal to US$1.00 per day, and that each family uses an average of 80 liters per day (roughly 15–20 lpcd) requiring four person trips to a source each day. Any improvement must be measured against a baseline; typically the existing condition without any intervention. In this project very limited data exists regarding the availability and location of sources, improved or otherwise, prior to the intervention. In the absence of that information, sources were iteratively added and randomly situated in the region of analysis using uniform distributions along the Cartesian coordinates. With seven initial sources, the average distance to a source and time spent collecting water are roughly equivalent to the values reported for this region (Hall *et al.* 2014).

MCC's experience in Mozambique is that the cost of a borehole, including project management, social mobilization and training, drilling, construction, and installation is roughly $20,000. In this case, the project life was assumed to be 20 years and the analysis assumes that recurring benefit streams and costs have a similar duration; although the different handpump components may have different useful lives.

For each alternative the following factors are calculated: (1) the number of proposed point sources to be constructed; (2) the aggregate annual cost of operating and maintaining those point sources; (3) the corresponding price that has to be charged to collect revenue equal to the O&M cost where 30 Meticais is roughly equal to one US dollar; (4) the fraction of households willing to pay that fee; (5) the results of the p-median model which is the population-weighted distance between households and sources – this is the distance cost if all households used the nearest water source; (6) the weighted cost which allows that some households are willing to pay to use their nearest source while others are not and use the source upon which they traditionally relied; (7) the daily distance savings, which is the difference between the base case (no new sources) and the current alternative; (8) the corresponding annual pecuniary value of the distance savings; (9–11) the PVs of the distance costs, O&M, and capital costs; and (12) the NPV. The Annex (available online at http://www.iwaponline.com/washdev/005/128.pdf) provides a table based on the values presented above.

RESULTS AND DISCUSSION

From a review of each of the alternatives, planners can better understand how various scenarios compare against each other across several metrics that may be important to them – average distance to a source, total annual cost, or required user fees – and to understand key tradeoffs between these variables. An important feature of the p-median location model is that while each additional source results in a reduction in the average distance between households and sources the incremental change is decreasing. At the same time, the addition of new sources results in higher fees and lower willingness to pay. Figure 3 demonstrates an important element of this model – namely the impact of willingness to pay on the average distance to a source. If households are asked to pay a higher tariff to ensure the financial sustainability of the handpump, then some households may elect not to pay and therefore not to use an improved source; presumably these households continue to use their traditional sources. The distance savings, and ultimately benefits, for any scenario is the difference between the base case (zero new sources) and the alternative in question with *n* new sources. There is a point at which the benefits associated with incremental distance reductions are outweighed by the number of additional households who elect not to pay and thus travel longer distances. Even if access to a handpump is provided at no cost, there will still exist a point at which the incremental distance benefits are outweighed by the costs of installation.

A second observation is that with the values used in this conceptual case, the distance-related benefits are vastly greater than the capital and operating costs associated with the project. Not surprisingly, Figure 4 shows that NPV has a similar maximum to distance savings. In this illustrative example, the optimal number of point sources is seven. At this service level, the corresponding fee should be set to 12 Meticais per month per household and 88% of households would pay this amount. The avoided one-way distance is 4,811 km per day.

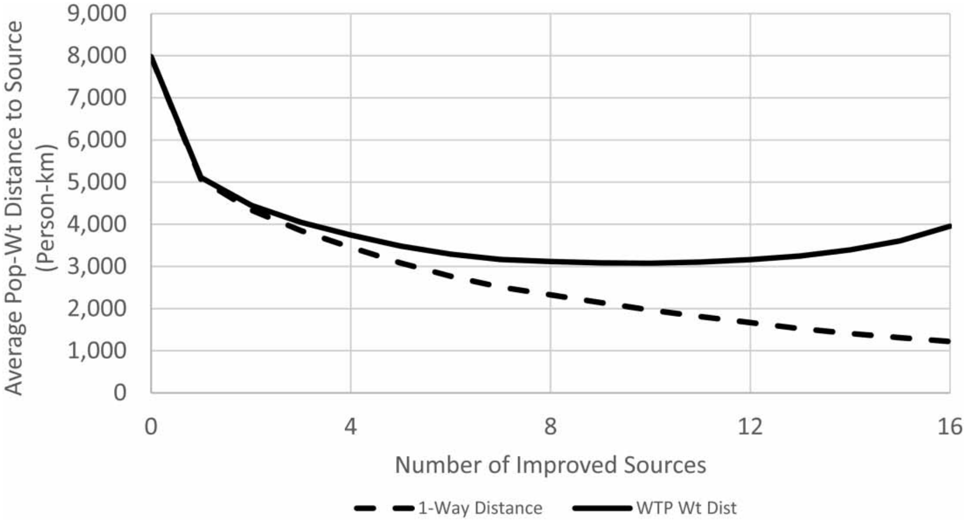

Figure 3 | Population-weighted distances between households and sources.

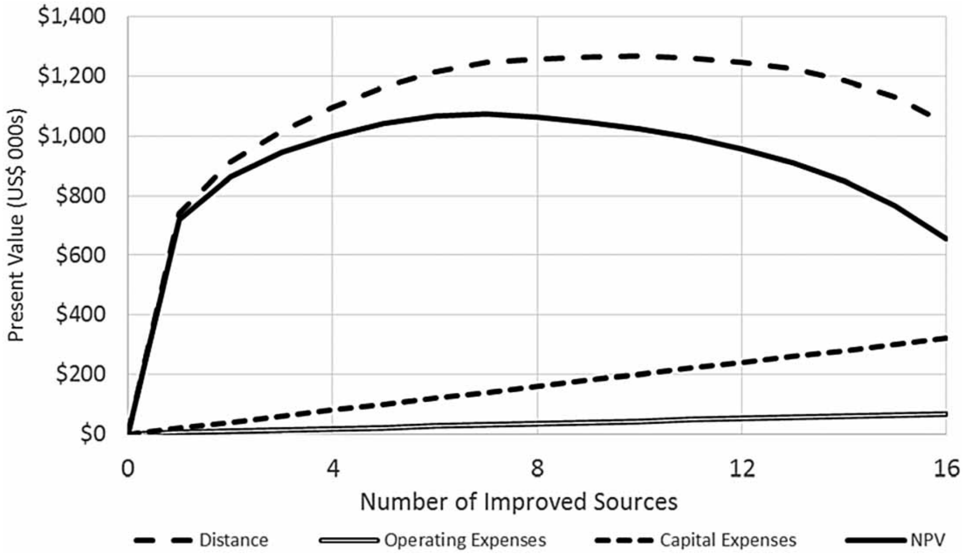

Figure 4 | Benefits, costs, and NPV by number of sources.

CONCLUSION

This paper makes some important contributions to the literature on the provision of rural water services in developing countries. First it presents a new conceptual model that seeks to maximize the distance-related benefits by using mathematical models and readily available information to help identify optimal solutions. Second, whereas many rural water projects focus on neighborhoods and use national services standards as the design standard, this paper presents an alternative approach based on planning at a regional level. Because households are in fact traveling long distances, planning at a larger scale may be more appropriate. Third, to the extent that households are being asked to make a financial payment to use the system it suggests that there will be a breakeven point at which the incremental benefit of reduced distances is balanced by the increasing number of households who elect not to use the system on the basis of cost.

A few words of caution are warranted. First, the focus of this model is to identify the optimal number of improved sources. As a result, the principal interest is in the relative rates of change of the principal components of the costs and benefits and not in their absolute values per se. Recall that the distance benefits, and specifically the changes in the distance benefits, are assumed to greatly exceed all other benefits; this implies that these other benefits can be treated as if they are essentially constant and that it is reasonable to ignore their effects on the optimal number of point sources. Nonetheless, for the range of values used in this paper, the NPV is always positive, often strongly so.

Second, the willingness to pay model represents a cross-section of consumer preferences in a fixed point in time. In fact, infrastructure projects of the type under consideration typically have long design lives – on the order of 20 years. Hence in the design of rural water supply systems, and particularly in determining the number and location of point sources, planners ought to consider the temporal planning dimension. Frequently, the presence of new infrastructure may serve as a catalyst for increased demand: 'Build it and they will come,' which essentially expresses skepticism about the predictive capabilities of willingness to pay models. Clearly, any effective long-term strategy for improving water supply infrastructure must be able to accommodate the inevitable changes in demand that occur over time. Hence, planning water supply systems, rural or otherwise, must be considered a dynamic process and planning strategies must be continuously reevaluated at the start of each planning cycle.

Third, the household willingness to pay for access to an improved water source in this model is principally a function of price whereas in fact it is probable that it is a function of both price and the distances to the traditional sources and the proposed distance to a new source; likewise in areas with substantial changes in topology, relative elevations may also have an impact on willingness to pay. Further while linear representations of demand may be suitable for small segments of the demand curve they may not correctly reflect demand over a larger range.

Fourth, the model does not allow for seasonal variations. It is often observed that households revert to closer, less salubrious, sources once they become available with the onset of the rainy season. This may lead to substantial reductions in distance- and time-related benefits and the ability to collect revenue from users.

Fifth, the model identifies the number and location of the optimal number of point sources based on a single criterion (the highest NPV). In reality any decision about the optimal number and location of point sources is a function of multiple objectives, only one of which was illustrated in this paper. Many others are un-modeled and might include equity considerations, political constraints, and so on. Therefore, for all these reasons, the model presented here is not intended as a substitute for judgment but rather as an aid to decision makers.

This model provides planners with the kinds of information needed to design a rural water supply system, including the number and location of point sources, types of fees that need to be charged, and how many users will patronize each source. Similarly, the addition of a sensitivity analysis could shed light on prioritization strategies for data collection. This model also provides planners with a simple tool for re-evaluating alternatives using updated information. The p-median model can be easily modified to require that certain sites must, or must not, have a source; this allows the analyst to deal with circumstances when it might infeasible to locate a source in a particular neighborhood (say for hydrogeological reasons) or if a source had to be located at a particular location (because of the presence of a health clinic). The method is simple and can be run on a modest desktop or laptop computer allowing for designers and planners to rapidly develop and examine various planning alternatives and compare them to a benchmark solution.

From this model it becomes clear that the design of rural water supply projects requires information on demand for these services, the actual cost of provision of the services and particularly those that will be passed on to consumers, geospatial and demographic data on communities, and the location of existing sources (improved or otherwise). With respect to implementing such a model, it is not the intent of this paper to suggest that optimal planning solutions can be determined without substantial input from planners or consumers or the need to undertake geophysical studies to identify exact locations. On the contrary, these models hold out the potential of facilitating the decision making process by highlighting the perils of traditional, standards-based planning while presenting an alternative approach that identifies a band of near optimal solutions and the range of associated prices that must be charged, the corresponding number of point sources, and their locations. The key point is that district level officials need to be actively involved and engaged in the planning process.

DISCLAIMER

The author is an employee of the MCC. The opinions expressed in this paper are his and do not necessarily reflect the view of MCC or the US Government. The author declares that other than the compensation he receives as an MCC employee, he has no financial interest in the research described in this paper.

REFERENCES

Burke, R. 2014 Needs Assessment for MCC's Rural Water Service Delivery Projects: MCC Internship Report. College Park, MD.

Carl Bro International. 1997 ESRDF Handbook for Rural Water Supply and Sanitation.

Carter, R., Chilton, J., Danert, K. & Olschewski, A. 2014 *Siting of Drilled Water Wells: A Guide for Project Managers*. RWSN, St Gallen, Switzerland. Retrieved from http://www.rural-water-supply.net/en/resources/details/187.

ESRI. 2012 Desktop Help 10.0 - Location-allocation analysis. Retrieved 02 February 2015. http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//004700000050000000.

Giné-Garriga, R., de Palencia, A. J.-F. & Pérez-Foguet, A. 2013 Water-sanitation-hygiene mapping: an improved approach for data collection at local level. *The Science of the Total Environment* **463–464**, 700–711.

Governments of the United States, Mozambique 2007 *Compact between MCC and Mozambique*, Washington, DC. Retrieved 20 April 2015. https://www.mcc.gov/documents/agreements/compact-mozambique.pdf.

Hall, R. P., Davis, J., van Houweling, E., Vance, E. A., Carzolio, M., Seiss, M. & Russel, K. 2014 *Impact Evaluation of the Mozambique Rural Water Supply Activity*. Blacksburg, VA. Retrieved 20 April 2015. http://www.researchgate.net/publication/266375992_Impact_Evaluation_of_the_Mozambique_Rural_Water_Supply_Activity_Acknowledgements.

Ho, J. C., Russel, K. C. & Davis, J. 2014 The challenge of global water access monitoring: evaluating straight-line distance versus self-reported travel time among rural households in Mozambique. *Journal of Water and Health* **12** (1), 173–183.

Hopkins, O. S., Lauria, D. T. & Kolb, A. 2004 Demand-based planning of rural water systems in developing countries. *Journal of Water Resources Planning and Management* **130** (1), 44–52.

Hutton, G. & Haller, L. 2004 *Evaluation of the Costs and Benefits of Water and Sanitation Improvements at the Global Level*. WHO, Geneva. Retrieved 20 April 2015. http://www.who.int/water_sanitation_health/wsh0404.pdf.

Instituto Nacional de Estatistica 2014 Data Portal: Mozambique. Retrieved 13 April 2014. www.ine.gov.mz.

Jiménez, A. & Pérez-Foguet, A. 2011 Water point mapping for the analysis of rural water supply plans: case study from Tanzania. *Journal of Water Resources Planning and Management* **137** (5), 439–447.

Millennium Challenge Corporation (MCC) 2013a *Chapter 16: Guidelines for Countries Proposing Water and Sanitation=Projects*. Retrieved 02 February 2015. http://www.mcc.gov/pages/docs/doc/compact-development-guidance-chapter-16.

MCC. 2013b Chapter 5: Guidelines for Economic and Beneficiary Analysis. Retrieved 02 February 2015. http://www.mcc.gov/pages/docs/doc/guidelines-for-economic-and-beneficiary-analysis.

Ministry of Water, Livestock Development 1997 *District Operational Manual*. Dar es Salam, Tanzania.

Nauges, C. & Whittington, D. 2009 Estimation of water demand in developing countries: an overview. *The World Bank Research Observer* **25** (2), 263–294.

Pearce, J. & Howman, C. 2012 RWSN Water Point Mapping Group: A Synthesis of Experiences and Lessons Discussed in 2012 (p. 28). RWSN, St Gallen, Switzerland. Retrieved 20 April 2015. http://www.rural-water-supply.net/_ressources/documents/default/1-450-3-1357715729.pdf.

ReVelle, C. S. & Swain, R. W. 1970 Central facilities location. *Geographical Analysis* **2** (1), 30–42.

RWSN Executive Steering Committee. 2010 *Myths of the Rural Water Supply Sector: RWSN Perspective No 4*. RWSN, St Gallen, Switzerland. Retrieved 20 April 2015. http://www.rural-water-supply.net/en/resources/details/226.

Van Wijk-Sijbesma, C. A. & Smet J. E. M. 2002 Planning and management. In: Small Community Water Supplies: Technology, People and Partnership. IRC International Water and Sanitation Centre, Delft, The Netherlands. Retrieved 20 April 2015. http://www.ircwash.org/resources/small-community-water-supplies-technology-people-and-partnership.

WashCost 2012 10 key messages about WASH costs and service levels in Mozambique.

WaterAid 2010 *Water Point Mapper*. WaterAid, London.

Wedgwood, A. & Sansom, K. 2002 *Willingness to Pay Surveys: A Streamlined Approach*. WEDC, Loughborough, UK. Retrieved 20 April 2015. http://wedc.lboro.ac.uk/resources/books/Willingness-to-pay_Surveys_-_Complete.pdf.

Welle, K. 2010 *Strategic Review of WaterAid's Water Point Mapping in East Africa*. London, UK. Retrieved 24 April 2015. http://www.wateraid.org/~/media/Publications/east-africa-water-point-mapping.pdf.

First received 4 June 2014; accepted in revised form 6 April 2015. Available online 13 May 2015