Centre for Efficiency and Productivity Analysis

Working Paper Series
No. WP07/2013

Using Fieldwork, GIS and DEA to Guide Management of Urban Stream Health

H. K. Millington, J. E. Lovell and C. A. K. Lovell

Date:
November 2013

School of Economics
University of Queensland
St. Lucia, Qld. 4072
Australia

ISSN No. 1932 - 4398

Using Fieldwork, GIS and DEA to Guide Management of Urban Stream Health

H. K. Millington,[a] J. E. Lovell and C. A. K. Lovell[b,*]

[a] Australian Rivers Institute, Griffith University, Nathan QLD 4111, Australia
[b] CEPA, School of Economics, University of Queensland, Brisbane QLD 4072, Australia

Abstract

Aquatic ecosystems are vulnerable to threats from human activity. Urban stream ecosystems are especially vulnerable to urbanisation of surrounding land use, and interest continues to grow in improving the health of urban streams. We study 30 sites along two highly urbanised streams in Brisbane, Australia. Field research generated a suite of stream health indicators at each site. Spatially explicit geographic information system (GIS) techniques were used to determine metrics of nearby land use that put stress on stream health at each site. Population density is also considered as a stressor. Data Envelopment Analysis (DEA) is applied to individual health indicators (one at a time) and multiple stress indicators to construct a suite of best-practice frontiers, from which ecological efficiencies and response elasticities are calculated at each site. DEA is also used to aggregate stream health indicators into a stream health index for each site, and to aggregate land use stress indicators into a land use stress index for each site. A second round of DEA is then applied to the stream health index and multiple stress indicators, and a third round of DEA is then applied to the stream health and land-use stress indices to create an overall ecological performance index for urban streams (EPIUS). Empirical findings show significant deviations beneath best practice, wide variation in response elasticities, and numerous dominance relationships in all three exercises. Each of these findings can provide guidance to those responsible for allocating scarce resources in an effort to improve the management of the health of Brisbane's urban streams.

Keywords: aquatic ecosystem health, field research, GIS, DEA, ecological efficiency

**Corresponding author*. Tel.: +61 7 3346 7057; fax: +61 7 3365 6666.
E-mail address: k.lovell@uq.edu.au (C. A. K. Lovell)

Using Fieldwork, GIS and DEA to Guide Management of Urban Stream Health

1. Introduction

Many local catchment groups and local governments invest time and money into protecting and rehabilitating urban streams and urban riparian zones. It is therefore important from the outset to know what constitutes a realistic target and to identify which areas will be most responsive so that scarce resources can be allocated for maximum benefit. In this paper we develop an analytical framework that incorporates measures of both urban stream health and land use stress. The framework identifies sites most in need of attention and generates response elasticities of alternative stream health measures to changes in alternative land use stress measures. We provide an empirical application to two highly urbanized streams in Brisbane, Australia to illustrate how the framework achieves these objectives.

Addressing stream health and stream stress factors

Due to the complex nature of urban stream ecosystem processes, the mechanisms by which land cover and hydrological alteration impact urban stream health have not been directly demonstrated although correlations have been illustrated. Ecosystem health impacts on urban streams, such as altered hydrology and channel morphology, habitat fragmentation and loss, high nutrient levels, pollutants, and invasive species of plants and animals have been collectively referred to as the “urban stream syndrome” (Allan 2004, Meyer et al. 2005, Walsh et al. 2005b). However in different locations the relative importance of these urban stream stressors varies. For example, in-stream connectivity was found to be very important to fish assemblage, pollution levels and habitat quality in Puerto Rico (Ramírez et al. 2012), hydrological alteration associated with levels of catchment-scale impervious surface was found to be the most important land-cover feature impacting macroinvertebrate and fish community structure in Victoria, Australia (Walsh et al. 2005a) and in Georgia, USA (Roy 2004), and intact riparian tree cover at the reach scale was found to have a detectable benefit on macroinvertebrate community structure in Victoria (Thompson and Parkinson 2011).

The most common approach to planning and prioritising stream rehabilitation projects is based on ‘available land opportunities’, with the result that most stream rehabilitation activities are undertaken in headwaters and small tributaries, although the habitat and land use changes which are most severe are commonly in lowland floodplains and deltas (Bernhardt et al. 2005, Hermoso et al. 2012). Inspired by the systematic conservation planning used for reserve design (Ardron et al. 2010, Margules and Pressey 2000), a systematic planning approach for river rehabilitation has been recommended (Hermoso et al. 2012). This “efficient planning” approach proposed by Hermoso et al. (2012) allows for the efficient selection of areas for rehabilitation based on socio-economic constraints and facilitates decision-making by integrating and prioritising the trade-offs among multiple rehabilitation actions using multiple objective optimisation (Czyzak and Jaszkiewicz 1998). Systematic planning has not been applied in many rehabilitation projects (Hermoso et al. 2012), and the few examples that exist (see Llewellyn et al. 2005, Peters and Marmorek 2001, Steel 2008) have not used multiple-objective optimisation methods.

Ecosystem processes at the whole-catchment scale are fundamental to systematic rehabilitation planning. However the scale of planning and the scale of implementation of rehabilitation measures do not have to be the same (Hermoso et al. 2012), and in light of evidence of the different scales of key drivers of impact on urban stream health in different areas and under different climatic conditions (Walsh et al. 2005a, Thompson and Parkinson 2011, Engman and Ramírez 2012), tools to help identify the appropriate scales to target management intervention are required. Failure to apply river management intervention at a scale appropriate to capture the driving processes has been blamed for the poor performance of many rehabilitation activities (Hermoso et al. 2012, Bernhardt et al. 2005).

Proponents of systematic planning have not yet articulated an analytical framework of how such an approach would be applied to river rehabilitation. Indeed, there is a need for a framework which allows easier integration and comparison of alternative rehabilitation actions for managers to consider and which address the driving ecosystem processes. The aim of the present study is to further elucidate options for protection and rehabilitation of freshwater urban ecosystems as well as the scale of mitigation efforts that might be required. It is possible that at some locations in the current study the freshwater urban ecosystem might not be a priority to protect for its own biotic integrity due to the high levels of stress it might be under.

The south east Queensland approach

Healthy Waterways is a not-for-profit, non-government organisation devoted to the protection and improvement of waterways in South East Queensland (SEQ), Australia (www.healthywaterways.org). It operates an Ecosystem Health Monitoring Program (EHMP) that monitors whether the health of regional waterways is improving or deteriorating. It uses a broad range of biological, physical and chemical indicators of ecosystem health, including fish and invertebrate biodiversity metrics, ecosystem process metrics and water quality metrics. The EHMP indicators have been selected to capture reach, local and catchment-scale processes.

The EHMP was fully implemented in 2002/03. 135 freshwater stream sites in SEQ are sampled twice annually (in spring and autumn), their health indicators are measured, and report cards are made public in annual reports. The overall health of a site is measured relative to an agreed reference condition (Bunn et al. 2010). The EHMP sites include relatively pristine, forested sites as well as rural sites and a small number of urban sites. Many local councils use the results of the EHMP as a guide to how well they are protecting their creeks. However, depending on the level of urbanisation or rural land cover, returning streams to good health may not be feasible. For example, it may require returning a substantial percentage (approximately 80%) of hydrologically-active land to mid-dense forested cover to return streams to “healthy” grades (Report Card Grade of B+ to A) (Sheldon et al. 2012). This may not be realistic in many locations where there is extensive pre-existing infrastructure. The EHMP does not have tools to aid prioritisation of sites for protection or rehabilitation based on the level of stress they are under.

The general poor health of urban streams in the Lower Brisbane Catchment is well documented. In the 2013 EHMP report, the Lower Brisbane Catchment, which includes the Norman Creek and Bulimba Creek sub-catchments, received a grade of

D-, down from a grade of D+ in 2012 but up from a grade of F in the previous six years (www.healthywaterways.org). Grades are based exclusively on stream health indicators, and although individual EHMP health indicators help identify the most likely stressors, and EHMP acknowledges "significant signs of stress," particularly at urban sites, EHMP does not consider stressors in the calculation of report card grades.

Effective management of urban stream health requires an understanding of the interrelationships among health indicators, among stressors, and between the two. Achieving such an understanding requires an analytical framework that incorporates both health indicators and stressors. We introduce such a framework below, and we use it throughout the paper.

Evaluating the relative performance of stream sites

Data Envelopment Analysis (DEA) is an analytic benchmarking technique developed by Charnes et al. (1978). Its linear programming structure makes it an accessible and easily understood management tool. Though the most common applications are to measure the performance of public agencies and private firms, it can be applied to myriad problems that seek to investigate the relative efficiency of transforming inputs to outputs. This ecological performance analysis of 30 stream sites within the Bulimba and Norman Creek sub-catchments of the Brisbane River is the first application of DEA to urban stream ecology.

A DEA requires "inputs" and "outputs" and a maximising or minimising objective. In this urban stream application it is assumed that the objective of decision makers is to maximise stream health given the stress levels at specific sites. Therefore stream health indicators derived from field research are considered the "outputs" to be maximised, and low stress indicators calculated in GIS are the "inputs". Relatively efficient sites are able to achieve better stream health conditional on their low stress inputs. A DEA is particularly useful when comparing "like with like", and sites in the Norman Creek and Bulimba Creek sub-catchments have relatively homogeneous environmental features (climate, topography, soils, geology and natural vegetation), and being contained in the Lower Brisbane Catchment, they can all be classed as degraded.

In contrast to ordinary least squares regression analysis, in which an estimated *function* intersects the data, DEA envelops the data with an estimated *frontier* that determines "best practice", rather than average practice. DEA sheds light on the underlying cause-effect relationships by comparing the ability of similar sites to withstand the stresses caused by an urban environment. When the stressors are included in the performance analysis, it is possible to measure how "ecologically efficient" a site is in achieving its level of health. The ecological performance of an inefficient site can be measured by its distance from the best-practice frontier. This measure of ecological inefficiency helps to better understand why some sites' ecosystem health measures appear to respond better to the surrounding land use and provides a realistic target for stream restoration and protection.

A commendable feature of DEA is the ability to combine or aggregate multiple inputs and multiple outputs that are measured in their own units. This is an important simplifying feature when handling the diverse variety of data required in a study of ecological performance.

It is important to emphasise two features of DEA: (1) it accounts for *both* stream health *and* stress factors when evaluating sites; and (2) its evaluation of each site is *relative* to the performance of all other sites in the sample, rather than to an agreed reference condition. Thus, while the Bulimba and Norman creeks may well be in generally poor health, a DEA can distinguish degrees of poor health at the sampled sites, and it can relate degrees of poor health to causal stress indicators at sampled sites.

DEA also can complement systematic planning by assisting in both adaptive management of rehabilitation projects already implemented (Hermoso et al., 2012, Wenger et al., 2009), and proactive management to identify which catchments are priorities for future rehabilitation.

The paper unfolds as follows. In Section 2 we discuss the analytical framework we use to relate stream health to the stressors that threaten stream health. The framework begins with fieldwork, continues with GIS, and concludes with DEA. We describe our data and the processes by which they have been generated in Section 3. In Section 4 we discuss our empirical findings. Section 5 concludes.

2. Methods

2.1 Fieldwork

Brisbane, the state capital located on the Lower Brisbane River, is the major population centre in SEQ with approximately 2.2 million people in the greater Brisbane area, while the population in the region is approximately 3.1 million and growing at 2.9% per annum in an area of 22,600 km^2 (Australian Bureau of Statistics 2013, Abal et al. 2005). Population growth in SEQ continues to be one of the key threats to the sustainability of stream health in the region. SEQ is a sub-tropical region with rainfall patterns that are influenced to some degree by the summer cyclones but also frequently by the northward extension of temperate weather systems. The wet season from October to March delivers approximately twice as much rain as the relatively dry season from April to September (Australian Bureau of Meteorology 2013). The city of Brisbane has a mean annual rainfall of 1450 mm.

Stream health data (macroinvertebrate and water quality) were collected for 30 sites in Bulimba Creek and Norman Creek (BCNC) sub-catchments, southern tributaries of the Lower Brisbane River during the post-wet season, April 2010. Sites were selected from within the BCNC catchments to include as great a range as possible of catchment impervious land cover and associated stormwater drains and piping, as well as the greatest range possible in tree, grass and riparian land cover at the reach and catchment scales. Sites were also designed to include nested sites and longitudinally connected sites that covered an extensive component of the catchments. However site selection was constrained by site accessibility and therefore sites that were completely piped underground were not included for obvious access reasons.

This focus study of BCNC aims to spatially investigate the role of the riparian zone in urban stream health by investigating spatial-scale (reach, local and catchment scale) impacts of tree and total vegetation (tree and grass) cover, impervious surface and stormwater piping. Because the BCNC study is interested in examining in-stream, riparian and catchment processes, it is important to include indicators reflective of these processes which might indicate differences due to stressors at different scales. Macroinvertebrate indicators have been associated with reach and catchment-scale riparian condition (Bunn et al. 2010) as well as broader catchment processes (Walsh et al. 2005b), dissolved oxygen and temperature have been more strongly associated with reach and catchment-scale riparian condition than broader catchment processes, while pH and conductivity have usually been associated most strongly with broader catchment processes (Bunn et al. 2010).

Therefore, the selected suite of health indicators (a subset of the indicators that form the basis of the EHMP (Bunn et al. 2010)) includes: (1) water quality data (pH, conductivity, dissolved oxygen diel readings and temperature diel readings); and (2) macroinvertebrates. The macroinvertebrate data were analysed for SIGNAL2 score (Chessman 2003) ("order-class-phylum" level), EPT (Lenat 1988), and family richness (Resh et al. 1995), however only SIGNAL2 was used in subsequent analysis because EPT and family richness had too many zero values and an insufficient range of values to be useful in the spatial scale analysis. SIGNAL2 is a rapid assessment procedure that accounts for pollution grades for each family or order-class-phylum level.

2.2 GIS

A novel feature of this study is the combination of DEA with spatially explicit geographic information system (GIS) land-cover metrics tailored to the stressors affecting urban stream ecosystem health. GIS techniques were used to determine land-cover metrics for different land-cover configurations and population density to test which of these variables had the most explanatory power for each of the stream ecosystem health variables. ArcGIS 9.3 and 10.0 were used for metric generation. The key spatial scales considered were reach, local and catchment scales but consideration of land-cover configuration via different spatially-explicit metrics such as inverse distance-weighted (IDW) metrics and areal buffer metrics as well as non-spatial metrics (lumped, i.e. each pixel is weighted equally) facilitated further investigation of the dominant processes impacting urban stream health. Custom python scripts were developed to generate and automate calculation of these land-cover metrics (Millington 2013).

In urban catchments such as these, many streams are piped underground or flow under roads and are thus not simply connected on the surface. Hence, in order to apply contemporary digital elevation model (DEM)-based hydrologic modeling techniques to create a connected stream network that incorporates both surface and piped flow, it was necessary to construct an artifical surface stream network by treating the piped flow as surface flow. This was achieved by burning (imprinting) existing stream and pipe networks into the DEM (Millington 2013). The methods for generating the artificial stream network generally follow those of Gironás et al. (2010), in which piping, streams and other known drainage are burnt into the DEM, with streams burnt in a greater distance than elements of the anthropogenic drainage.

The raw raster land-cover data set initially comprised nine land-cover classes which were grouped into three broad land-cover types: impervious surface, standing woody vegetation (trees), and a combined vegetation metric (grass and trees) (Millington 2013). Based on the land-cover data, impervious surface was considered to comprise layer codes 5 (roads/asphalt), 6 (buildings/structures), and 7 (concrete). Tree covered surface included layer code 4 (trees), and vegetation was considered to include layers 3 (grass) and 4. Water and bare ground/rock were not included in the land-cover calculations in this study as they could not easily be categorised as tree, vegetation or impervious surface. The categories of tree and impervious surface were required because the study was designed to detect relative importance of reach-scale riparian zone and catchment-scale impervious surface. Vegetation was included as a further variable because it constituted a large surface area and also to consider whether any detected benefit of tree cover was due to shading and other processes associated with tree cover or simply the presence of pervious surface.

2.3 DEA

DEA is a linear programming technique developed to compare how efficient decision-making units (DMUs) are at achieving key objectives, generally to either maximise outputs obtained from a certain level of inputs, or minimise inputs required to produce a certain amount of outputs. DEA envelops the data with an estimated *frontier* that determines the “best practice” and provides a benchmark against which the relative efficiency of DMUs is measured.[1]

Non-marketed goods and services

Output prices are an important component of many economic performance evaluations and aggregation techniques. DEA was originally proposed for use in the public sector, where outputs are typically non-marketed and output prices are typically missing. Examples include education, health care, emergency and corrective services, etc. In the first empirical application of DEA, Charnes et al. (1981) were concerned with evaluating the efficiency of a U.S. public school education program (“program efficiency”), how it was managed (“management efficiency”), and with distinguishing the two. Over 30 years on, DEA programs and the vast majority of empirical applications remain independent of prices, whether or not they exist.

Environmental and ecological problems

Environmental and ecological problems fit into the missing prices framework; environmental and ecological variables are not priced on markets. Early applications of DEA to environmental issues appeared over 20 years ago (Färe et al. 1989, 1993). These and many subsequent studies were directed primarily at two closely related problems: (1) how best to include undesirable by-products of the production of desirable outputs (e.g., air pollutants resulting from electricity generation, or methane gases from livestock) into a producer performance evaluation model; and (2) how to measure the shadow prices (or marginal abatement costs) of the undesirable outputs. Marginal abatement costs are the cost of abating a marginal amount of an undesirable output (e.g., the cost of a marginal reduction of CO2 emissions or of catchment impervious surfaces) without reducing any desirable output or increasing any input. The measurement of shadow prices is potentially very useful, since shadow prices can be used in place of missing output prices. These models have been evaluated by Scheel (2001), and by Dykhoff and Allen (2001), who specifically mention “ecologically motivated applications of DEA.” Kuosmanen and Kortelainen (2005) provided a more recent evaluation, using DEA to construct an “encompassing eco-efficiency index” as the ratio of an economic value added index to an environmental damage index created by aggregating environmental pressures, with an application to road transport (which generates environmental pressure from CO2, acidification, smog, particle dispersion and noise).

Among the many applications of DEA to environmental and ecological problems, a few are particularly relevant to the present study. Each study measures environmental-economic performance by incorporating environmental indicators into an economic framework. Some also measure the opportunity cost of sub-optimal performance, and some measure shadow prices of environmental variables.

Hof et al. (2004) used DEA to relate 12 inputs (undesirable forest and rangeland condition indicators such as habitat disturbance and toxic chemical releases) to 13 outputs (human activity indicators such as timber harvest, livestock grazing and outdoor recreation) across 3,011 counties in the USA. Their objective was to identify areas where there was maximum potential for improving forest and rangeland condition; "areas that are currently under the most stress but could be made better simply by managing resource use more efficiently…because environmental impacts (as reflected in our condition indicators) could be reduced without the economic and social costs associated with reducing economic activity." They concluded that large-scale improvements in environmental condition across many indicators were unlikely without a reduction in human activity.

Macpherson et al. (2010) followed Hof et al. by using a variant of DEA to conduct an "environmental-economic" evaluation of 134 watersheds in the USA. Their data base contained four inputs (watershed characteristics including per cent impervious surface, road density and percent of stream-length with adjacent agricultural riparian land cover), four desirable outputs (a mix of socio-economic and environmental variables including percent wetland and percent interior forest) and six undesirable outputs (environmental problems including pollution and exotic aquatic and terrestrial species). They examined four models, each containing different combinations of desirable outputs reflecting different management objectives. All four models fit the data very well, with efficiency scores clustering closely beneath unity, indicating little potential for improvement and making discrimination among alternative management objectives difficult.[2]

2.3.1 DEA Best-Practice Frontier

In most production contexts inputs and outputs are positively correlated, in the sense that more of any input can produce more of any output. In this stream study, good stream health is positively correlated with low levels of stress, but not all conventional stream health indicators (outputs) are indicators of *good* health, and not all conventional stress indicators (inputs) are indicators of *low* stress. Consequently, in order to maintain the presumed positive relationship between inputs and outputs, some conventional stream health indicators are converted to good health indicators and some conventional stress indicators are converted to low stress indicators, prior to DEA analysis. We describe the data and the transformations in Section 3.[3]

A stream health maximising DEA model

A health maximising problem with r=1,…,s good health indicators y_r, i=1,…,m low stress indicators x_i, and j=1,…,N DMUs, which in this study are stream sites, can be stated mathematically as the program:

$$Max \quad \phi_o$$

Subject to:

$$\sum_{j=1}^{N} \alpha_j x_{ij} \le x_{io} \qquad i = 1,\ldots,m$$

$$\sum_{j=1}^{N} \alpha_j y_{rj} \ge \phi_o y_{ro} \qquad r = 1,\ldots,s \qquad \text{[M.1]}$$

$$\sum_{j=1}^{N} \alpha_j = 1, \alpha_j \ge 0 \qquad j = 1,\ldots,\mathrm{N}$$

In the primal envelopment program [M.1] the performance of site o is being evaluated. x_{ij} denotes the level of the ith low stress indicator observed at site j and y_{rj} denotes the level of the rth good health indicator observed at site j. α_j is the weight assigned to site j in the evaluation of the performance of site o, and $\sum_{j=1}^{N} \alpha_j = 1$ is a convexity constraint that permits varying returns to scale (RTS) at different site sizes, and allows for the possibility that a given percent increase in the low stress indicators might generate a larger or smaller percent increase in the stream health indicators. The performance of site o is evaluated relative to a convex combination of adjacent best-practice sites. These best-practice sites have similar levels of stress and have positive values of α, while all other sites have different levels of stress and have zero values of α. The optimal value of $\phi_o \ge 1$ indicates the potential increase in, or the scope to improve, all health indicators at site o without exceeding best practice and without increasing any low stress indicator. By convention, the health-oriented efficiency of site o is the *reciprocal* of the optimal value of ϕ_o, and so $0 \le \phi_o^{-1} \le 1$.

In addition to efficiency scores, response elasticities can be calculated from the input and output weights of the solution to the dual multiplier program:

$$Min \; \sum_{i=1}^{m} \delta_i x_{io} + \delta_o$$

Subject to

$$\sum_{r=1}^{s} \gamma_r y_{rj} = 1$$

$$-\sum_{r=1}^{s} \gamma_r y_{rj} + \sum_{i=1}^{m} \delta_i x_{ij} + \delta_o \ge 0 \qquad j = 1,\ldots,N \qquad \text{[M.2]}$$

$$\delta_i, \gamma_r \ge 0, \delta_o \; free$$

The non-negative weights attached to each indicator in the dual solution associated with [M.2] are multipliers, or normalised shadow prices. They measure the implicit value of each indicator, and the ratio δ_i/γ_r measures the impact on good health indicator y_r of a marginal change in low stress indicator x_i. Shadow price ratios are units-dependent, and difficult to compare across indicator pairs. However a ratio can be converted to a response elasticity, which is independent of units of measurement and easy to compare across indicator pairs, by means of

$$\varepsilon_{ri} = (\partial y_r/\partial x_i)\times(x_i/\phi y_r) = (\delta_i/\gamma_r)\times(x_i/\phi y_r).$$

The s×m response elasticities of each good health indicator with respect to each low stress indicator are evaluated at an efficient projection $(x_i,\phi y_r)$. The sign of the free parameter δ_o indicates the nature of scale economies at the efficient projection $(x_o,\phi y_o)$, with $\delta_o < 0$ indicating increasing RTS and $\delta_o > 0$ indicating decreasing RTS, also evaluated at an efficient projection. Response elasticities signal the proportional impact of a single low stress indicator on a single good health indicator, while the magnitude of scale economies signals the proportional impact of all low stress indicators on all good health indicators.

Dominance analysis[4]

A site (y_o,x_o) is said to dominate another site (y_n,x_n) if it has at least as much of each good health indicator and no more of each low stress indicator, i.e., if $y_o \geq y_n$ and $x_o \leq x_n$. A site that dominates another site is at least as healthy despite facing at least as much stress. Efficient sites need not dominate any other sites; they can be efficient merely by being different. Conversely, sites need not be efficient to dominate other sites; all they have to do is be at least as healthy and face at least as much stress. The value of dominance analysis is that a site that dominates many sites can provide managerially useful guidance on dealing with stress, and a site dominated by many sites has many role models from which to learn. Constructing a list of dominators for each site is easily handled with spreadsheet software.

2.3.2 DEA and Index Generation

Individual indicators (e.g. of stream health) provide independent, potentially conflicting, information, and it is useful to be able to aggregate individual indicators into a single index. A crucial issue is how diverse indicators are to be aggregated to construct an index. This involves two considerations. The first involves the choice of functional form of the aggregator function. There are many to choose from, and the choice is independent of the research area. The second involves the choice of weights to be applied to each indicator to reflect its relative importance. Most economic indices use market prices to weight economic indicators, but most ecological-environmental indicators lack market prices, so an alternative weighting scheme must be determined in order to construct an ecological-environmental index.

The aggregation of multiple ecological indicators into a scalar-valued holistic ecological index (or composite indicator) has attracted widespread attention, and some of the more relevant to the present study are listed below.

The Environmental Performance Index (EPI) (Emerson et al. 2012) has since its inception tracked the environmental performance of nations. Environmental performance is measured with an index obtained by aggregating numerous diverse environmental indicators. The EPI uses arbitrary weights informed by expert judgement, a procedure followed in much of the larger composite indicators literature. Zhou et al. (2006a), citing a lack of expert judgement, weighted three pollutants (SO_2, NO_2 and MP_{10}) equally in constructing an air quality index for

Chinese cities. In a pair of studies Hajkowicz (2006, 2007) used a variant of multi-criteria analysis to generate weights with which to aggregate environmental indicators in the Great Barrier Reef region into a water-service index, and to conduct a hypothetical reallocation of natural resource management funds across regions in Queensland. In the former study indicator weights were based on “stakeholder” (residents and visitors) preferences, and in the latter study indicator weights were provided by “decision makers” (commonwealth, state and regional governing bodies). It is worth stressing that in each of these examples weights are determined exogenously, prior to and independently of the performance evaluation exercise.[5]

There is an alternative to the use of predetermined exogenous weights provided by expert judgement, equality or stakeholder preferences in the development of aggregate indices. Models have been developed that generate endogenous weights as components of solutions to constrained optimisation problems. Linear programming is a popular way of formulating a constrained optimisation problem, and most variants of DEA are linear programs.

Tran et al. (2008) formulated a pair of linear programs that simultaneously generate endogenous weights with which to aggregate indicators and to rank units on the basis of the aggregate index. They applied this linear programming technique to the evaluation of the environmental performance of 141 US watersheds based on an aggregate index constructed from 50 indicators.

DEA can be used to aggregate indicators to construct an overall index (output, input and performance indices), and it offers some advantages over alternative methods. One advantage is that it constructs a best-practice frontier, based on indicators or indices, against which to evaluate the performance of each unit in the sample. A second is the ability to calculate response elasticities of the output indicators or index with respect to the input indicators or index. However for our purposes the most important advantage of DEA is that it does not require predetermined exogenous weights with which to aggregate individual indicators in the construction of an aggregate index; instead endogenously generated shadow prices emerge as weights from the dual multiplier program [M.2] above. These endogenously generated weights maximise the efficiency score of the unit under evaluation in comparison with the other units in the sample. In our context, although these weights may not reflect the consensus (if one exists) of ecologists on the relative importance of specific health and stress indicators, they do reflect the underlying processes that relate stress indicators to health indicators in our sample. A site’s endogenously determined weight profile reveals its relative strengths (relatively high weights) and weaknesses (negligible or zero weights) and thereby provides additional information to managers.[6]

Bellenger and Herlihy (2009) used a variant of DEA to construct a pair of environmental indices that account for chemical stress and ecological response at 130 stream sites in a single eco-region of the USA, with streams classified as minimally disturbed, moderately disturbed and heavily disturbed. The first index aggregates six macroinvertebrate metrics (included in the Environmental Protection Agency’s multimetric indicator MMI) into a stream health index. The second index aggregates the six macroinvertebrate metrics and three chemical stressors originating from non-point sources into an environmental performance index. Both indicator weights and performance indices vary across disturbance class. Later Bellenger and Herlihy (2010) used a stochastic envelopment technique similar to

DEA to evaluate the health of 215 streams in the same eco-region. The primary contribution of this study is its use of shadow prices, or “marginal performance estimates,” analogous to the multipliers that are obtained from the dual multiplier program [M.2] above, to weight each of the six macroinvertebrate metrics by their observed contribution to environmental performance. The objective of each study was index construction rather than efficiency measurement, because the identity of the drivers of environmental performance and the nature of the relationship between the two is “largely unknown.”

Azad and Ancev (2010) used a variant of DEA to measure the environmental performance of irrigated enterprises in 17 natural resource management regions (a total of 125 observations) within the Murray-Darling Basin of Australia. They constructed quantity indices for two inputs (volume of water applied and all costs excluding the cost of water), a single desirable output (gross revenue), and two undesirable outputs (an ecologically weighted water withdrawal indicator and a salinity impact indicator). They defined an environmental performance index as the ratio of the desirable to undesirable output quantity indices. The main finding was a wide variation in environmental performance across crops and across resource management regions.

Using DEA to construct an ecological performance index

The analysis of a single stream health indicator and multiple low stress indicators can be repeated many times over, using different stream health indicators and different sets of multiple low stress indicators, to search for common themes. A methodologically preferred alternative is to use DEA to aggregate stream health indicators into a single overall index of stream health at each site, and to aggregate low stress indicators into a single overall index of low stress at each site. This procedure combines diverse information provided by alternative health indicators into a health index, and combines diverse information provided by alternative low stress indicators into a single low stress index, thereby obviating the need to choose from among two groups of indicators. After the two indices have been constructed, a DEA analysis can be applied to the two indices to evaluate the overall ecological performance at each site.

A stream health index

Following Lovell (1995) and Lovell & Pastor (1999), a stream health index is obtained as the solution to the health maximisation program:

$$Max \quad \phi_o$$

Subject to:

$$\sum_{j=1}^{N} \alpha_j y_{rj} \geq \phi_o y_{ro} \qquad r = 1,\ldots,s \qquad [M.3]$$

$$\sum_{j=1}^{N} \alpha_j = 1, \alpha_j \geq 0$$

Program [M.3] aggregates s health indicators into a single health index. Its structure is identical to that of program [M.1] apart from the deletion of the m low stress constraints. Low stress indicators are ignored in this exclusively health-oriented program designed to aggregate multiple stream health indicators into a single stream health index. As in program [M.1], optimum $\phi_o \geq 1$ and the stream health index $SHI_o = Y_o(y_{1o,\ldots,}y_{so}) = \phi_o^{-1} \leq 1$. Site health index values range downward from 1 (the healthiest sites in the sample). The stream health index is analogous to the stream health report cards produced by EHMP (although it is constructed very differently) and the MMI index of Bellenger and Herlihy (2009).

We specify a DEA stream health maximisation program based on a single good health index and m low stress indicators as

$$Max \quad \phi_o$$

Subject to:

$$\begin{aligned}
&\sum_{j=1}^{N} \alpha_j x_{ij} \leq x_{io} && i = 1,\ldots,m \\
&\sum_{j=1}^{N} \alpha_j Y_j \geq \phi_o Y_o && \\
&\sum_{j=1}^{N} \alpha_j = 1, \alpha_j \geq 0 && j = 1,\ldots,\mathrm{N}
\end{aligned} \qquad [M.4]$$

Program [M.4] is used to estimate the relative efficiencies of sites, and its dual is used to estimate response elasticites. Program [M.4] is identical to program [M.1] apart from the replacement of s good health indicators with a single good health index. As in program [M.1], optimum $\phi_o \geq 1$ and the health index-based efficiency is $0 \leq \phi_o^{-1} \leq 1$.

A low stress index

The aggregation of m low stress indicators into a single low stress index is an input minimisation problem. The objective is to minimise overall low stress at each site, regardless of the stream health at each site. The DEA program for this input-oriented problem is

$$Min \quad \theta_o$$

Subject to:

$$\begin{aligned}
&\theta_o x_{io} - \sum_{j=1}^{N} \beta_j x_{ij} \geq 0 && i = 1,\ldots,m \\
&\sum_{j=1}^{N} \beta_j = 1, \beta_j \geq 0 &&
\end{aligned} \qquad [M.5]$$

Program [M.5] aggregates m low stress indicators into a single low stress index. Health indicators are ignored in this program designed to aggregate multiple low stress indicators into a single low stress index. Optimum $0 \leq \theta_o \leq 1$, and the low stress index is defined as $LSI_o = X_o(x_{1o},\ldots,x_{mo}) = \theta_o^{-1} \geq 1$, so that site low stress index values range upward from 1, the most stressed sites in the sample, with increases in θ_o^{-1} signaling reductions in stress. The low stress index is similar to the environmental damage index of Kuosmanen and Kortelainen (2005) and the ecologically weighted water withdrawal index of Azad and Ancev (2010).

An overall ecological performance index

Once a good health index and a low stress index have been calculated, DEA can be used again to construct a holistic measure of the performance of each site. DEA programs [M.1] and [M.2] are used again, with s+m indicators y_o and x_o replaced with two indices $SHI_o = Y_o(y_{1o,\ldots},y_{so})$ and $LSI_o = X_o(x_{1o},\ldots,x_{mo})$ created in programs [M.3] and [M.5]. The overall ecological performance index for urban streams (EPIUS) is the solution to the DEA program

$$Max \quad \phi_o$$

Subject to:

$$\begin{aligned} &\sum_{j=1}^{N} \alpha_j X_j \leq X_o \\ &\sum_{j=1}^{N} \alpha_j Y_j \geq \phi_o Y_o \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{[M.6]} \\ &\sum_{j=1}^{N} \alpha_j = 1, \alpha_j \geq 0 \qquad\qquad j = 1,\ldots,\mathrm{N} \end{aligned}$$

The optimal value of $\phi_o \geq 1$ indicates the potential increase in, or scope to improve, the good health index at site o without exceeding best practice and without increasing the low stress index. The health-oriented performance of site o is $0 \leq \phi_o^{-1} \leq 1$, and so the performance index $PI_o(X_o,Y_o)$ of sites ranges downward from 1.

3. Data

A sample consisting of 30 sites imposes a degrees of freedom constraint, which in turn requires a minimally informative variable list. The first screen was the explanatory power that a stream stressor variable exhibited for stream health metrics (Millington 2013). The second screen was variability across sites; if a variable exhibits little variation across sites, it provides little information and is excluded. A stream health metric was included in the DEA only if it exhibited a significant amount of variation that could be explained by the land-cover metrics. The third screen was based on the objectives of this study; it was desirable for the stream health indicators to retain a macroinvertebrate indicator and a water quality indicator, and for the land-cover and land-use indicators to retain a metric to represent each of three spatial scales, reach, local and catchment. On the basis of these screens, the variable list can be reduced to a manageable size.

From a preliminary analysis of nine stream health metrics and 20 land-cover/land-use variables, a set of *a priori* models was formulated for each stream health indicator. Candidate explanatory variables for the models were selected based on the student t-test ($p \leq 0.2$). The generalised least squares (GLS) function was used to fit the models and maximum likelihood (ML) was used to estimate the parameters. The model with the lowest Akaike's Information Criterion (AIC) statistic in the set was regarded as the "best" model, and if no one model clearly stood out as the best, inferences were based on the model-averaged parameter estimates. These statistical procedures guide selection of the most important explanatory metrics for stream health.

The two stream health indicators selected are a macroinvertebrate metric (SIGNAL2, an index that assigns a score to aquatic invertebrate families based on their tolerance for/sensitivity to pollution, with scores ranging from 1 for the most tolerant to 10 for the most sensitive) and a physical/chemical metric (temprange, water temperature range in Celsius). Additional physical/chemical stream health metrics exhibited insufficient variation across sites (maximum temperature), failed to respond significantly to variation in the selected land-cover metrics (conductivity, dissolved oxygen) or remained within healthy bounds at all sites (hydrogen concentration, pH).

The three land-cover/land-use metrics selected are the distance-weighted extent of impervious land cover (Eucdis, impervious land cover weighted by inverse Euclidean distance to the site), population density (popden, average population density from the 2006 Census in the upstream catchment of the site), and the extent of tree cover at the reach scale (treerip, percent vegetative cover provided by trees in the upstream riparian buffer 30 meters either side of the stream). popden is an important metric at the catchment scale for both SIGNAL2 and temprange, Eucdis is the only important metric at the local scale for SIGNAL2, and treerip is an important metric at the reach scale for both SIGNAL2 and temprange. It should be noted that popden may be acting as a surrogate for other land-cover metrics such as those which capture the extent of stormwater piping. But because it is the catchment-scale metric with the most support in the data it is likely that it also capturing other "largely unknown" explanatory factors associated with increased population. Popden and Eucdis can be considered as analogous to the socio-economic constraints mentioned in the systematic planning literature (e.g., Hermoso et al. 2012).

The need to define indicators of *good* health requires the inversion of temperature range to an indicator of temperature stability $(\text{temprange})^{-1}$ because a smaller stream temperature range enhances the growth, metabolism, reproduction and dispersal of aquatic organisms.[7] A similar need to define indicators of *low* stress requires the inversion of population density to population sparsity $(\text{popden})^{-1}$ and the extent of impervious land cover to $(\text{Eucdis})^{-1}$ because the greater the distance of impervious land cover from a stream the greater is the scope for attenuation before reaching the stream. Pollutants can be attenuated in the soil and hydrology impacts of stormwater flows are reduced.

Summary statistics for the two good health indicators and the health index, and the three low stress indicators and the low stress index, appear in Table 1. The two good health indicators have satisfactory inter-quartile ratios of approximately 1.5, and two of the three low stress indicators have larger inter-quartile ratios. Although $(\text{popden})^{-1}$ has a smaller inter-quartile ratio, it has a very large max/min ratio. The two indices inherit their inter-quartile ratios from those of their respective component indicators.

Insert Table 1 About Here

4. Findings and Discussion

4.1 DEA of the Primary Stream Health Indicator

The output-oriented DEA program [M.1] can be applied to each individual health indicator separately with three low stress indicators, or to two health indicators with three low stress indicators. We have conducted all three exercises, and to illustrate the insights DEA can provide while conserving space, we report the results of a DEA on the primary health indicator, SIGNAL2, and three low stress indicators.

The results appear in Table 2, which reports the output of program [M.1] with s=1, m=3, and which contains efficiency scores, scope to improve and response elasticities for each site. Additional results appear in Table 3, which reports dominance information for each site. Together these two Tables illustrate the wide range of information a DEA can provide decision makers.

The first four columns in Table 2 contain the data for each site. Column (1) contains the primary health indicator SIGNAL2, and columns (2)-(4) contain the low stress indicators $(\text{Eucdis})^{-1}$, $(\text{popden})^{-1}$ and treerip. Columns (5) and (6) contain the efficiency score ϕ^{-1} and the scope to improve ϕ for each site. Columns (7)-(9) contain the response elasticities of SIGNAL2 with respect to each of the three low stress indicators. Sites are ranked according to their efficiency scores.[8]

Insert Table 2 About Here

Table 2 reports a wide variation in relative efficiency for SIGNAL2, ranging from 1.0 (six sites) down to 0.39; the sites in most need of attention are those sites with the lowest efficiencies in the shaded region. Sites S_08 and S_25 have the two lowest efficiency scores of 0.39 and 0.43 respectively, and if well managed are capable of more than doubling their SIGNAL2 scores to 2.59 x 1.67 = 4.33 and 2.31 x 1.94 = 4.48 respectively, with no reduction in the stresses they face. Their low efficiency scores result from their poor health despite their average stress levels. At the other extreme, sites S_07, S_12 and S_13 are all best-practice sites, yet their health and stress levels differ enormously. S_07 is the healthiest site, and it is also one of the least stressed sites. In contrast, S_12 and S_13 are relatively unhealthy sites, with SIGNAL2 scores barely into the second quartile, and both are highly stressed sites.

Table 2 also reports wide variation in the response elasticities of SIGNAL2 with respect to the three low stress indicators. On average, the primary stream health indicator is slightly more responsive to marginal change in population density than to marginal changes in impervious surface or riparian tree cover, although all three response elasticities have mean values less than one. The low stress indicator to which stream health is most responsive at each site is the indicator with the largest response elasticity. Densely populated site S_22 illustrates. The high elasticity (6.63) with respect to $(\text{popden})^{-1}$ suggests that if it were possible to reduce population density within its watershed, the site would see a huge improvement in its stream health. Conversely, and more usefully, this large response elasticity provides a warning of the potentially deleterious effect of an increase in population density. The low stress indicator to which stream health is least responsive at each site is the indicator with the smallest response elasticity, and Table 2 contains many zero response elasticities. Sites S_15, S_16 and S_07 illustrate. Sites S_15 and S_16 have relatively little impervious surface and relatively large riparian tree cover, and so have zero response elasticities with respect to marginal improvements in either. Both have average population densities, and positive, albeit small, response elasticities with respect to marginal changes in population density. Site S_07 has the opposite features, and informs management that the primary health indicator responds positively to both a reduction in impervious surface and an increase in riparian tree cover. If management is interested in rehabilitating this site by increasing its SIGNAL2 score, resources would be best allocated to increasing riparian tree cover or to reducing impervious surfaces, a difficult task in an established urban landscape, but some mitigation steps via making pervious surfaces more effective and the use of water sensitive urban design (WSUD) may be possible.

A comment on the pattern of zeros among the response elasticities is warranted. Just over half of the 90 calculated response elastitities are zero, and half the sites assign zero weights to two of three low stress indicators. Variability of weights, or response elasticities, is inherent in DEA, which allows each site to select weights that put it in the most favourable light. Nonetheless, some would impose weight restrictions on program [M.2] to force them to align more closely with expert judgement. Rogge (2012) provides an interesting application to the EPI, in which he reports response elasticities and performance scores with and without weight restrictions. We have several reasons for not imposing weight restrictions. First, weight restrictions are exogenous to the model, and we prefer endogenous weights generated by the data and the model. Second, although we have endeavoured to compare like with like, sites have widely varying characteristics that are reflected in their widely varying response elasticities, and zero response elasticities convey important information to those tasked with the responsibility of allocating scarce resources in an effort to rehabilitate sites. Finally, several weight restriction models exist, and there is no consensus on which is best.

Table 3 lists the dominators for each site based on the primary health indicator SIGNAL2. If the SIGNAL2 score of site i is at least as large as the SIGNAL2 score of site o and $(\text{Eucdis})^{-1}$, $(\text{popden})^{-1}$ and treerip of site i are no greater than those of site o, then site i (weakly) dominates site o. Sites are ranked according to their efficiency scores, as in Table 2.

Insert Table 3 About Here

The dominator summary in Table 3 shows that ten sites are undominated, but only six of these undominated sites are best practice sites, demonstrating that it is not necessary to be best practice to be a useful role model. Site S_04 is not efficient, but it is undominated and dominates six other sites. Site S_26 is efficient, and it dominates 14 other sites. It has barely 20% tree cover in the reach-scale 200m riparian buffer and it has a relatively large impervious surface area, yet its SIGNAL2 score is higher than sites with more tree cover and smaller impervious surface area, demonstrating that an efficient and dominating site need not be a particularly healthy site (seven sites have higher SIGNAL2 scores). For example, site S_08 has less impervious surface area and over 70% tree cover, yet its SIGNAL2 score of 1.67 compares poorly with S_26's score of 3.5. It is dominated by 12 sites and dominates none. On both an efficiency criterion and a dominance criterion, site S_26 is a star and site S_08 a laggard. On these criteria Tables 2 and 3 show that a few other sites qualify as stars and a few qualify as laggards, and also illustrate the heterogeneity among stars and among laggards.

4.2 DEA of Stream Health, Stream Stress and Stream Performance

We begin by replicating the analysis summarised in Table 2, replacing the primary good health indicator SIGNAL2 with the good health index SHI created by program [M.3]. This exercise applies program [M.4], and provides an evaluation of the performance of sites based on the three low stress indicators and a broader measure of good health. We then aggregate the three low stress indicators into a low stress index LSI created by program [M.5], and we re-evaluate the performance of sites based on a single good health index and a single low stress index using program [M.6] to construct an overall ecological performance index EPIUS.

A stream health index for stream sites

We have created a stream health index SHI using the output-oriented program [M.4], in which the low stress indicators for each site are set to unity and the two good health indicators SIGNAL2 and $(\text{temprange})^{-1}$ are aggregated to the stream health index SHI. Table 4 reports the results. The stream health index SHI and the three low stress indicators are listed in columns (1)-(4). Efficiency scores and scope to improve appear in columns (5) and (6). Response elasticities of SHI with respect to the three low stress indicators appear in columns (7)-(9). Sites are ranked according to their efficiency scores.

Efficiency scores in Table 4 are very similar to those in Table 2, with many of the same stars and many of the same laggards. The pattern of response elasticities in Table 4 is also very similar to that in Table 2. Both similarities suggest that the primary good health indicator SIGNAL2 tracks the holistic stream health index SHI very closely.

Insert Table 4 About Here

We consider the response elasticities in somewhat more detail. Starting with $(\text{Eucdis})^{-1}$, more than half of the sites would benefit from a reduction in impervious surfaces, and S_13 would benefit most. Continuing with $(\text{popden})^{-1}$, a third of the sites could benefit from a less densely populated watershed, with S_22 the biggest beneficiary. If the impacts associated with increased population density (which might include increased catchment scale impervious surface, increased stormwater piping infrastructure, increased water pollution associated with increased levels of nutrients and heavy metals, or other urbanization stressors) could be mitigated, this site would see an improvement in health. Finally, more than a third of the sites report positive elasticities with respect to treerip in the reach-scale riparian zone. Planting more trees in the 200m reach-scale riparian zone at these sites would improve their health. Tree planting within the reach scale would see the best results at S_24 because its reach-scale tree coverage is the second smallest in the sample. Conversely, the 200m reach-scale riparian zone of site S_16 is almost totally covered with trees, its response elasticity is zero, and yet it is an unhealthy low performing site.

Insert Figure 1 About Here

Figure 1 provides a spatial presentation of the variation in SHI. The Norman Creek catchment is the small area bordered in pink, and the Bulimba Creek catchment is the large area also bordered in pink. Both catchments drain to the Brisbane River estuary to the north. The darker the area the healthier it is, and most healthy areas are upstream. However there are instances of relatively unhealthy sites upstream of healthier sites (e.g., S_19, S_20). The mean health in Norman Creek (SHI = 0.54) is substantially lower than in Bulimba Creek (SHI = 0.70). The mean impervious surface area for the Norman Creek catchment (36%) is higher than for the Bulimba Creek catchment (26%). In addition, Norman Creek has extensive sections of the creek network piped, and does not have most of its main channel maintained as free flowing channel with a relatively natural riparian zone as Bulimba Creek does.

A low stress index for stream sites

We have created a low stress index LSI using the input-oriented program [M.5], in which the good health indicators for each site are set to unity and the three low stress indicators $(\text{Eucdis})^{-1}$, $(\text{popden})^{-1}$ and treerip are aggregated to the low stress index LSI.

Table 5 reports the results, with sites ranked according to their low stress index. The stream health index and the low stress index appear in columns (1) and (2), efficiency scores appear in column (3), the scale elasticity appears in column (4), and dominance information appears in columns (5) and (6). We have already discussed SHI. Values of LSI range from 3.26 and 3.27 for the two least stressed sites to 1 for the four most stressed sites. A strong association of low stress with good health predominates, although not all healthy sites enjoy low levels of stress (e.g., S_18), and not all highly stressed sites are in relatively poor health (e.g., S_26, an efficient site). These outliers should attract the attention of those responsible for the health of these two streams. The scale elasticities in column (4) are all less than unity, suggesting that a given reduction in stress is likely to generate a less than proportionate improvement in health. The dominance information in columns (5) and (6) is very similar to that in Table 3, again suggesting that the primary health indicator SIGNAL2 is a reliable proxy for the more inclusive health index SHI.

Insert Table 5 About Here

Insert Figure 2 About Here

Figure 2 illustrates the spatial distribution of the low stress index. The darker the area the lower the stress. Norman Creek is generally more stressed (LSI = 1.22) than Bulimba Creek (LSI = 1.53), which is consistent with our finding that it is less healthy. The eastern side of Bulimba Creek is generally least stressed, which is consistent with it being less developed. Low stress sites are generally found in “headwater” sections (i.e., sites further up the catchment), although some sites higher in the catchment are highly stressed (e.g., S_12, S_19).

An ecological performance index for stream sites

We have created an overall ecological performance index EPIUS from output-oriented program [M.6] using SHI and LSI. The results appear in column (3) of Table 5. Most high-performing sites are healthy, but S_26 has just average health despite facing the highest stress levels in the sample. S_26 also dominates 20 sites, making it a star. Conversely, most poor-performing sites are unhealthy, regardless of the stress they face. S_20 is a poor performer despite having nearly average stress, and is dominated by 13 sites, making it a laggard.

Insert Figure 3 About Here

The spatial distribution of EPIUS values is evident in Figure 3. The darker the area the higher is its performance. Best performing sites include S_26 (EPIUS=1), which is one of the most stressed sites (LSI=1) but is moderately healthy (SHI=0.74). S_07, S_10, S_17 and S_18 also have EPIUS=1. S_07, S_10 and S_17 are classed as headwater streams in this study because no other sites are sampled further upstream to them and they pick up relatively small catchment areas. S_07 and S_10 are the two least stressed sites in the study. S_18 is downstream of a site with a lower performance index (EPIUS=0.44).

Apart from S_26, Norman Creek sites as a whole generally have lower EPIUS scores (EPIUS = 0.57) than Bulimba Creek sites (EPIUS = 0.80), and one factor that a visual inspection of sites highlights is the amount of stream network fragmented and piped underground in Norman Creek. Star site S_26 appears to have a high level of surrounding grass and trees in the local area both upstream and downstream, although there are not a lot of riparian trees directly on the stream segment. The presence of local scale grass in the surrounding landscape (both upstream and downstream) should therefore be investigated as a possible way to improve health.

Health v. Performance

We noted in the introduction that Healthy Waterways operates EHMP, which monitors the health of SEQ waterways, using a broad range of biological, physical and chemical health indicators. EHMP does not incorporate stress indicators when it creates report cards. We have constructed three performance measures that do incorporate stress indicators. The first is based on SIGNAL2 and three low stress indicators, and is reported in Table 2. The second is based on SHI and three low stress indicators, and is reported in Table 4. The third is based on SHI and LSI, and is reported in Table 5. We also have calculated three rank correlation coefficients, between each health indicator or index and the corresponding performance measure. The objective is to determine whether ignoring the stresses confronting urban streams, as EHMP does, paints an accurate picture of their relative performance.

Based on the three rank correlation coefficients, the answer is "on average, yes," but not as significantly as one might expect, and with some large discrepancies. The rank correlation coefficients between health and performance measures in Tables 2, 4 and 5 are 0.74, 0.73 and 0.92, respectively. Stream health tracks stream performance fairly well, especially when low stress indicators are aggregated to a low stress index that smooths out some of the variability in the component indicators. To cite one example of discrepancy, site S_26 is an efficient site in all three exercises, yet its health, as measured by SIGNAL2 and by SHI, would warrant a grade of B- or C+ on its report card.

5. Conclusions

Brisbane City Council is interested in data supporting decision making for new developments that support ecological, environmental, economic and social benefits for Norman Creek (Brisbane City Council 2011, 2013).

Urban stream management

Urban streams are complex systems making it difficult to identify critical stressors and their relationships to urban stream health. There are multiple factors interacting at several scales and correlations and regression analysis alone are not always able to tease out the stressors and mechanisms by which they impact stream health. By comparing the relative ability of similar sites to withstand the stress caused by an urban environment DEA allows the researcher to study the relationships further.

DEA can help determine the existing state of the stream and what improvements are possible and realistic. The stream health index SHI identifies the healthiest streams when stress is not accounted for; it is similar to a report card on health. The low stress index LSI identifies the relative stress levels at each site; and helps understand why some sites are healthier than others. The overall performance index EPIUS identifies best-practice sites on the stream health frontier and quantifies the amount of improvement that could be achieved by inefficient sites given the existing spatial configuration of their urban landscape (proximity of impervious surfaces, reach scale riparian vegetative cover, broader catchment scale impacts indicated by population density). When the three indices are mapped using GIS software (Figures 1, 2 and 3), the visual presentation can further assist managers in their task of identifying the most suitable sites for rehabilitation.

The combination of both stressors and health indicators that can be included in a DEA offer promise as a tool for management of freshwater ecosystem health. Careful selection of the type and scale of the variables used is necessary for an ecologically meaningful outcome. This deterministic approach of DEA can be coupled with a statistical approach such as model fitting and model averaging to provide insights to management from these two perspectives to indicate realistic management goals. By exhibiting increased health with increased low stress, the results of the DEA analysis support the selection of the stream health and inverse stream stress variables used here; this evidence is consistent across Table 2, which uses stream health indicator SIGNAL2, and Table 4, which uses the stream health index SHI. This evidence also suggests that riparian rehabilitation would benefit urban stream ecosystem health at some sites in this study area. However in some cases the sites may be too stressed to benefit, and the expected outcomes need to be realistic.

The information in Table 5 raises several questions for management. How do some of the more highly stressed sites manage to be relatively healthy? Why are some of the low stressed sites relatively unhealthy? How should resources be allocated to make the optimal improvement in stream health? By observing the situation at best-practice sites and dominator sites, management can get an idea of what improvements are feasible for the less efficient sites. Elasticity measures for SHI with respect to the low stress index LSI help to identify the sites most likely to respond to changes in stressors and whether there is likely to be a proportionate change in health. In Table 5 all elasticities are less than one, suggesting that there are decreasing returns to scale; a proportional increase in stream health will be less than a proportional reduction in stress. However the elasticities for half the sites are close to 1.00 suggesting that in some cases, the changes could be equiproportionate. These are the sites that management may want to consider for rehabilitation.

Conversations with representatives from star and laggard sites also provides a useful way to consider what might be allowing one site with similar levels of stress to be healthier and to pose further hypotheses. Considering possible explanations led to further testable hypotheses - does the presence of terrestrial vegetation in the local area improve health of this site, and is there a way to measure the in-stream longitudinal potential connectivity of these stream sites. It became evident that in-stream longitudinal connectivity varied greatly among sites, and appeared to be especially low for Norman Creek sites, in which extensive sections of the stream network are buried or fragmented by concrete and stormwater piping. These questions have been addressed (e.g., Hughes et al. 2013) and continue to be investigated (Millington 2013).

Considering an EPIUS that accounted for both health and stress highlighted star sites which were healthier for a given level of stress than lower performing sites. These results prompted a desktop investigation into which important aspects of land-cover or land-use were unaccounted for in the BCNC study. Our measure of impervious surface is a metric that captures the impervious land cover near the site (distance-weighted Euclidean distance, but with a high weighting on the land cover close to the site). It might be beneficial to repeat the analysis including instead alternative land-cover metrics such as percent piped channel in the upstream "effective riparian buffer" 30m either side of stream that captures a more extensive catchment area and would potentially have more differentiation with the reach-scale tree cover (Millington 2013).

Another possible explanation for better performance was the amount of terrestrial vegetation in the landscape around sites. This suggests that surrounding terrestrial vegetation may be an important explanatory variable, prompting further GIS metric development for future studies (Millington 2013).

Some sampled metrics, such as dissolved oxygen and conductivity, were not included in the DEA because they responded positively to measures of increased urbanisation, which was unexpected from the literature. However the literature reports variability in how these metrics respond to urbanisation, especially in low flow situations. The stream health data used in this DEA were collected during one season, and as such the sample is a snapshot in time. Sites with very low flows may be unable to support as high a level of macroinvertebrate diversity as they could when flows are greater. But this assessment provides a starting point for considering local and catchment-wide stresses, health and priorities for rehabilitation for these streams, which could form part of *ad hoc* or systematic planning approaches.

Finally, it would be desirable to narrow the analytical and data gaps between our study and EHMP, which has been in place for over a decade. We already have noted the analytical gap between health and performance. It would be interesting to collect data on an expanded suite of health indicators used in EHMP reports for these sites or additional sites. The five categories of stream health measures covered by the EHMP were not covered in this study. Effectively only water quality and macroinvertebrates were. The EHMP study was only partially concerned with urban streams but some of their indicators are still relevant. There are also other metrics related to urban sites that are not collected by the EHMP such as heavy metals that may be available in other urban stream data sets such as those captured by BCC and other urban councils. Larger samples provide more information than smaller samples and this study was based on a small sample of 30 urban sites.

Variables	Labels	min	Q1	mean	Q3	max
Site Good Health Indicators						
SIGNAL2	y_1	1.38	2.19	2.84	3.48	4.81
temprange		0.7	1.53	2.19	2.4	6.7
$(\text{temprange})^{-1}$	y_2	0.15	0.42	0.56	0.66	1.43
Site Low Stress Indicators						
Eucdis		0.07	0.19	0.25	0.32	0.44
$(\text{Eucdis})^{-1}$	x_1	2.25	3.14	5.21	5.2	14.29
popden		2.61	14.07	15.17	17.2	21.83
$(\text{popden})^{-1}$	x_2	0.046	0.058	0.079	0.071	0.383
treerip	x_3	0.157	0.474	0.659	0.899	0.998
Site Health Index						
SHI	Y	0.37	0.52	0.67	0.84	1.00
Site Low Stress Index						
LSI	X	1.00	1.09	1.47	1.52	3.27

Table 1 Descriptive Statistics for the Data Set

Site ID	y_1 = SIGNAL2	x_1 = (Eucdis)$^{-1}$	x_2 = (popden)$^{-1}$	x_3 = treerip	ϕ^{-1}	Scope to improve	ε_{11}	ε_{12}	ε_{13}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
S_07	4.81	14.286	0.159	0.895	1.00	1.00	0.07	0.00	0.28
S_12	2.36	2.253	0.058	0.157	1.00	1.00	1.87	0.00	0.00
S_13	2.27	2.425	0.047	0.410	1.00	1.00	2.87	1.05	0.00
S_14	2.90	13.831	0.046	0.935	1.00	1.00	0.00	2.49	0.00
S_18	4.61	3.602	0.063	0.923	1.00	1.00	0.00	0.03	0.30
S_26	3.50	2.835	0.050	0.191	1.00	1.00	0.00	1.22	0.00
S_17	4.60	13.755	0.114	0.886	0.98	1.02	0.00	0.06	0.28
S_04	3.26	2.790	0.066	0.496	0.96	1.05	1.60	0.00	0.00
S_05	3.69	3.136	0.059	0.899	0.94	1.07	1.15	0.00	0.00
S_15	3.92	11.723	0.071	0.932	0.85	1.18	0.00	0.03	0.00
S_10	3.93	7.375	0.383	0.900	0.84	1.19	0.04	0.00	0.29
S_30	3.88	3.939	0.072	0.994	0.84	1.19	0.02	0.00	0.00
S_03	2.92	3.792	0.051	0.414	0.81	1.24	0.00	1.21	0.00
S_11	2.96	3.769	0.054	0.549	0.77	1.30	0.00	1.20	0.00
S_28	2.79	2.957	0.068	0.822	0.76	1.32	1.16	0.00	0.00
S_01	3.42	3.782	0.069	0.962	0.74	1.35	0.02	0.00	0.00
S_29	2.76	3.418	0.069	0.482	0.70	1.43	0.02	0.00	0.18
S_24	2.23	3.202	0.059	0.185	0.67	1.48	0.00	0.00	1.90
S_27	2.37	4.831	0.071	0.265	0.65	1.54	0.00	0.00	0.14
S_23	2.50	4.167	0.069	0.658	0.59	1.69	0.02	0.00	0.23
S_06	1.72	2.572	0.059	0.558	0.58	1.74	1.69	0.00	0.00
S_02	2.00	3.154	0.051	0.793	0.56	1.80	0.00	1.21	0.00
S_22	1.82	3.169	0.049	0.537	0.54	1.87	0.10	6.63	0.00
S_20	2.08	3.694	0.063	0.472	0.53	1.89	0.02	0.00	0.18
S_16	2.42	7.639	0.086	0.998	0.52	1.93	0.00	0.04	0.00
S_21	2.19	6.369	0.086	0.640	0.52	1.94	0.03	0.00	0.23
S_09	2.20	5.200	0.091	0.835	0.49	2.05	0.03	0.00	0.28
S_19	1.38	2.591	0.064	0.343	0.46	2.19	1.68	0.00	0.00
S_25	1.94	4.787	0.061	0.922	0.43	2.31	0.00	1.17	0.00
S_08	1.67	5.200	0.070	0.712	0.39	2.59	0.00	0.04	0.24
mean	2.84	5.208	0.079	0.659	0.74	1.48	0.41	0.55	0.15

Table 2 Efficiency Scores and Response Elasticities in the Primary Stream Health Indicator Model [M.1] and [M.2]

Site ID	ϕ^{-1}	Dominators											
S_07	1.00	-											
S_12	1.00	-											
S_13	1.00	-											
S_14	1.00	-											
S_18	1.00	-											
S_26	1.00	-											
S_17	0.98	-											
S_04	0.96	-											
S_05	0.94	-											
S_15	0.85	S_18											
S_10	0.84	-											
S_30	0.84	S_18											
S_03	0.81	S_26											
S_11	0.77	S_26											
S_28	0.76	S_04	S_26										
S_01	0.74	S_05	S_18	S_26									
S_29	0.70	S_26											
S_24	0.67	S_12											
S_27	0.65	S_26											
S_23	0.59	S_03	S_04	S_11	S_26	S_29							
S_06	0.58	S_12	S_13										
S_02	0.56	S_13	S_26										
S_22	0.54	S_13											
S_20	0.53	S_12	S_13	S_24	S_26								
S_16	0.52	S_01	S_03	S_04	S_05	S_11	S_18	S_23	S_26	S_28	S_29	S_30	
S_21	0.52	S_03	S_04	S_11	S_12	S_13	S_24	S_26	S_27	S_29			
S_09	0.49	S_03	S_04	S_11	S_12	S_13	S_23	S_24	S_26	S_27	S_28	S_29	
S_19	0.46	S_12											
S_25	0.43	S_02	S_03	S_05	S_11	S_12	S_13	S_24	S_26				
S_08	0.39	S_03	S_04	S_06	S_11	S_12	S_13	S_20	S_22	S_23	S_24	S_26	S_29

Table 3 Dominators in the Primary Stream Health Indicator Model [M.1] and [M.2]

Site ID	Y = SHI	x_1 = $(\text{Eucdis})^{-1}$	x_2 = $(\text{popden})^{-1}$	x_3 = treerip	ϕ^{-1}	Scope to improve	ε_{Y1}	ε_{Y2}	ε_{Y3}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
S_05	0.84	3.136	0.059	0.899	1.00	1.00	1.18	0.00	0.00
S_07	1.00	14.286	0.159	0.895	1.00	1.00	0.00	0.00	0.00
S_10	1.00	7.375	0.383	0.900	1.00	1.00	0.00	0.00	0.00
S_12	0.56	2.253	0.058	0.157	1.00	1.00	1.27	0.00	0.00
S_13	0.51	2.425	0.047	0.410	1.00	1.00	2.35	1.15	0.00
S_14	0.62	13.831	0.046	0.935	1.00	1.00	0.00	2.34	0.00
S_17	1.00	13.755	0.114	0.886	1.00	1.00	0.00	0.00	0.00
S_18	0.96	3.602	0.063	0.923	1.00	1.00	0.05	0.00	0.00
S_21	1.00	6.369	0.086	0.640	1.00	1.00	0.25	0.00	0.17
S_26	0.74	2.835	0.050	0.191	1.00	1.00	0.00	1.14	0.00
S_04	0.71	2.790	0.066	0.496	0.97	1.03	1.21	0.00	0.00
S_01	0.89	3.782	0.069	0.962	0.92	1.09	0.05	0.00	0.00
S_30	0.84	3.939	0.072	0.994	0.86	1.16	0.05	0.00	0.00
S_15	0.83	11.723	0.071	0.932	0.85	1.17	0.00	0.11	0.00
S_11	0.65	3.769	0.054	0.549	0.80	1.25	0.00	1.13	0.00
S_03	0.61	3.792	0.051	0.414	0.79	1.27	0.00	1.14	0.00
S_28	0.59	2.957	0.068	0.822	0.76	1.32	1.20	0.00	0.00
S_22	0.52	3.169	0.049	0.537	0.72	1.38	0.08	5.93	0.00
S_24	0.50	3.202	0.059	0.185	0.71	1.42	0.00	0.00	1.42
S_09	0.68	5.200	0.091	0.835	0.69	1.46	0.07	0.00	0.00
S_29	0.58	3.418	0.069	0.482	0.69	1.45	0.16	0.00	0.15
S_02	0.49	3.154	0.051	0.793	0.65	1.54	0.00	1.14	0.00
S_27	0.51	4.831	0.071	0.265	0.65	1.53	0.00	0.00	0.19
S_08	0.59	5.200	0.070	0.712	0.62	1.61	0.00	0.31	0.17
S_16	0.61	7.639	0.086	0.998	0.61	1.64	0.00	0.13	0.00
S_23	0.55	4.167	0.069	0.658	0.60	1.66	0.18	0.00	0.19
S_06	0.37	2.572	0.059	0.558	0.56	1.77	1.23	0.00	0.00
S_19	0.37	2.591	0.064	0.343	0.55	1.82	1.23	0.00	0.00
S_20	0.44	3.694	0.063	0.472	0.52	1.93	0.17	0.00	0.14
S_25	0.45	4.787	0.061	0.922	0.48	2.08	0.00	1.11	0.00
mean	0.67	5.208	0.079	0.659	0.80	1.32	0.36	0.52	0.08

Table 4 Efficiency Scores and Response Elasticities in Model [M.4]

Site ID	Y = SHI	X = LSI	ϕ^{-1} = EPIUS	RTS = ε_{YX}	# of DOMINATORS for this site	# of sites this site DOMINATES
	(1)	(2)	(3)	(4)	(5)	(6)
S_10	1.00	3.27	1.00	0.00	3	0
S_07	1.00	3.26	1.00	0.00	2	1
S_17	1.00	2.40	1.00	0.00	1	2
S_09	0.68	1.94	0.68	0.00	8	0
S_16	0.61	1.85	0.61	0.00	10	0
S_21	1.00	1.81	1.00	0.13	0	5
S_30	0.84	1.53	0.85	0.12	3	2
S_15	0.83	1.53	0.85	0.12	3	2
S_08	0.59	1.49	0.60	0.11	9	0
S_23	0.55	1.48	0.57	0.11	11	0
S_01	0.89	1.46	0.91	0.11	1	6
S_27	0.51	1.41	0.53	0.11	11	0
S_29	0.58	1.40	0.60	0.11	8	2
S_18	0.96	1.33	1.00	0.10	0	9
S_20	0.44	1.32	0.46	0.92	13	0
S_25	0.45	1.30	0.48	0.92	12	1
S_28	0.59	1.27	0.64	0.92	6	6
S_05	0.84	1.25	0.92	0.92	0	11
S_04	0.71	1.21	0.80	0.91	1	9
S_11	0.65	1.14	0.78	0.91	1	8
S_19	0.37	1.14	0.44	0.91	9	0
S_06	0.37	1.11	0.46	0.91	8	1
S_24	0.50	1.08	0.63	0.91	6	4
S_03	0.61	1.08	0.76	0.90	2	10
S_02	0.49	1.08	0.62	0.90	5	4
S_22	0.52	1.05	0.67	0.90	3	7
S_26	0.74	1.00	1.00	0.90	0	20
S_14	0.62	1.00	0.84	0.90	1	16
S_12	0.56	1.00	0.75	0.90	2	10
S_13	0.51	1.00	0.68	0.90	3	6
Mean	0.67	1.47	0.74	0.52	5	5

Table 5 Overall Ecological Performance EPIUS, Response Elasticities and Dominance in Index Model [M.6]

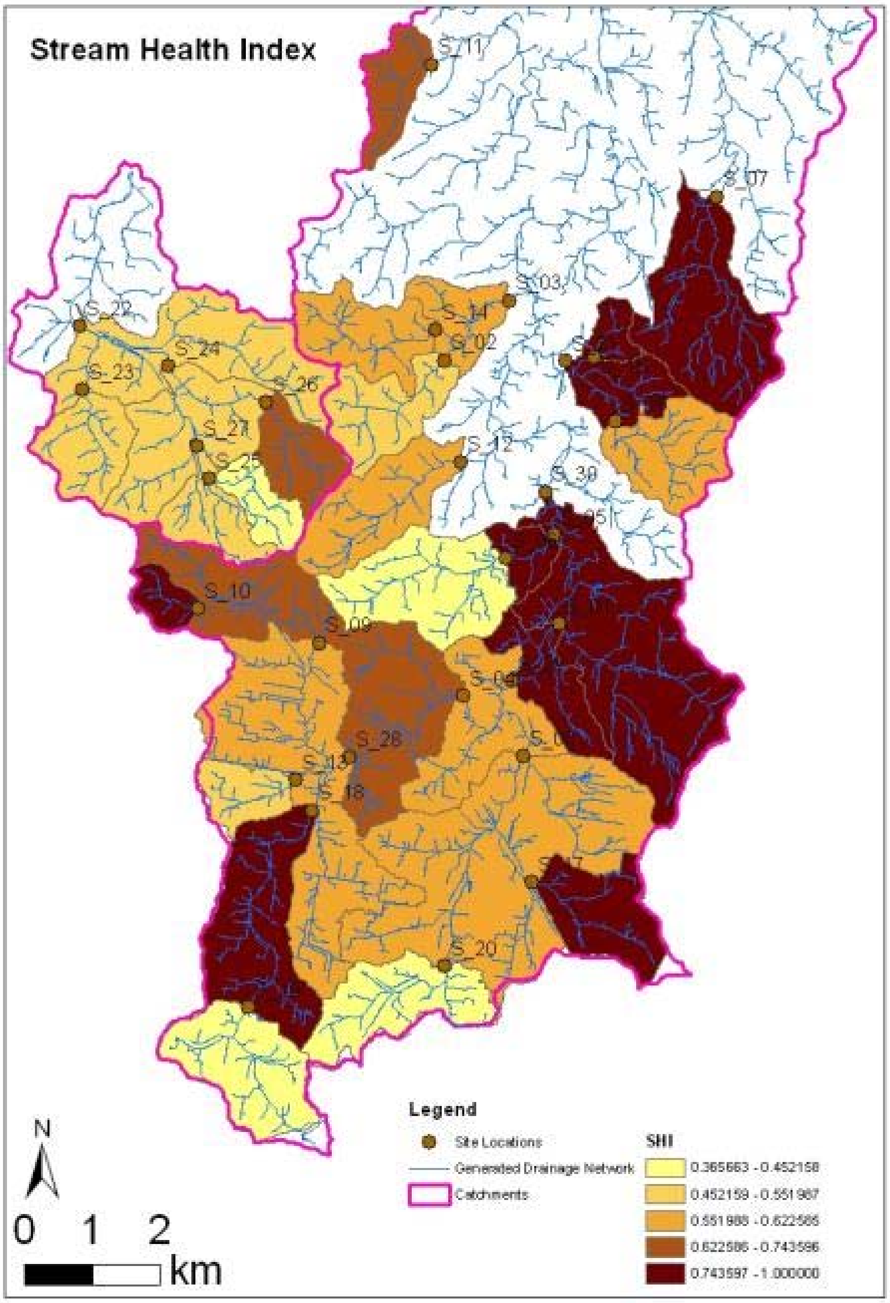

Figure 1 The Spatial Arrangement of the Stream Health Index SHI

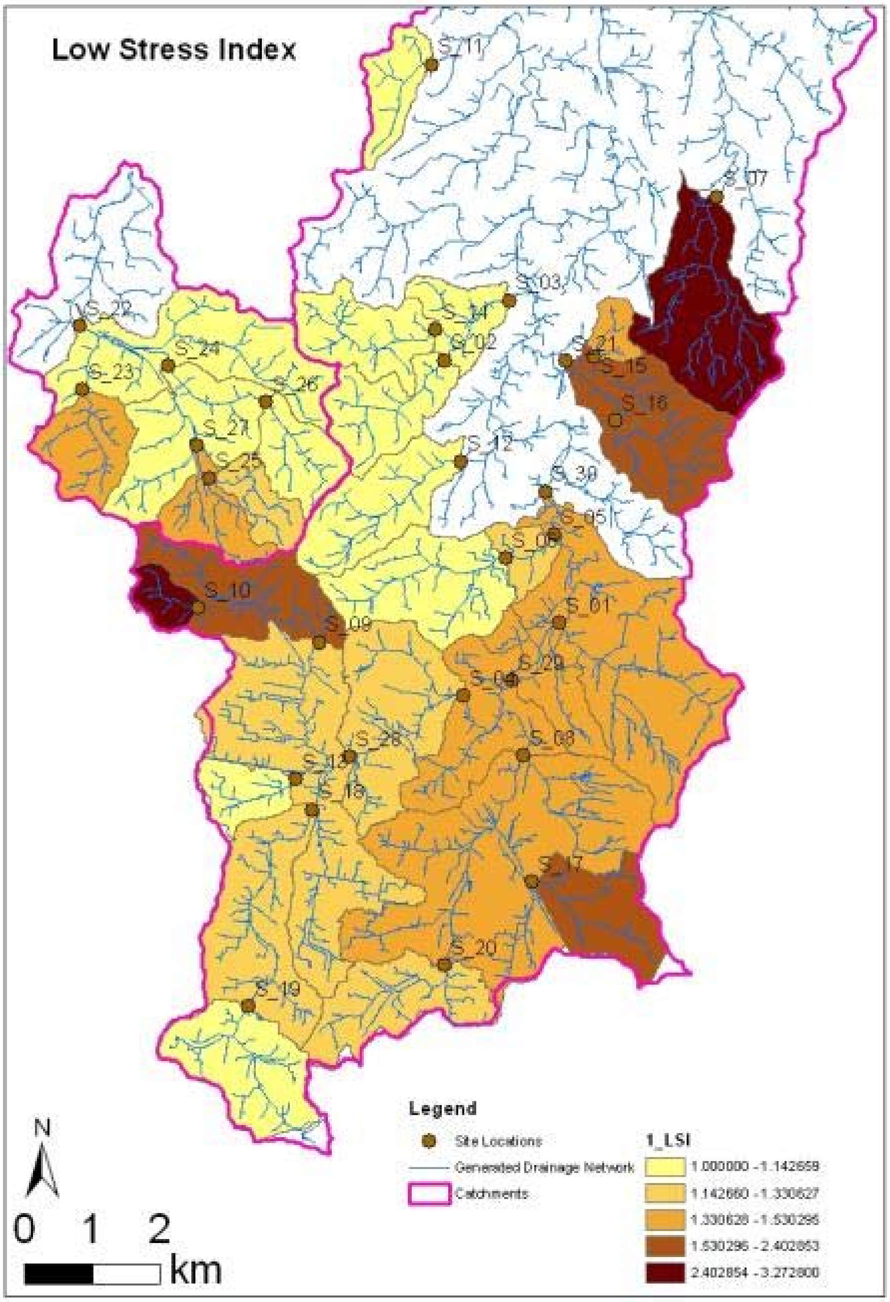

Figure 2 The Spatial Arrangement of the Low Stress Index

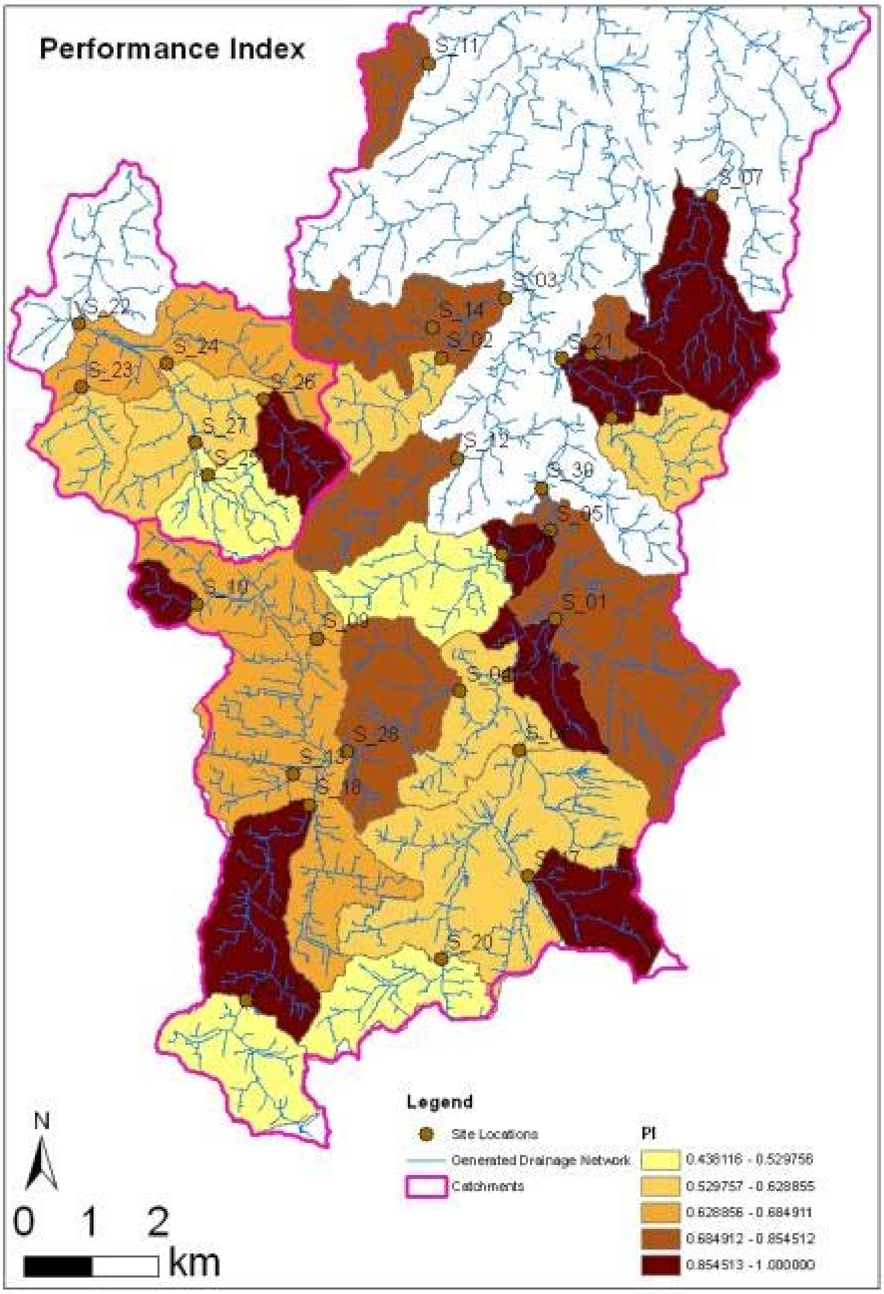

Figure 3 The Spatial Arrangement of the Overall Performance Index EPIUS

References

Abal, E. G., Bunn, S. E., Dennison, W. C. (Editors), 2005. Healthy Waterways, Healthy Catchments: Making the Connection in South East Queensland, Australia. www.healthywaterways.org

Allan, J. D., 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annu. Rev. Ecol. Evol. Syst. 35, 257-284.

Ardron, J. A., Possingham, H. P., Klein, C. J., Editors, 2010. Marxan Good Practices Handbook, Version 2. Pacific Marine Analysis and Research Association, Victoria, BC, Canada.

Australian Bureau of Meteorology, 2013. Climate statistics for Australian locations. www.bom.gov.au/climate/averages

Australian Bureau of Statistics, 2013. Regional Population Growth, Australia. www.abs.gov.au/ausstats/abs@.nsf/Previousproducts/3218.0

Azad, M. A. S., Ancev, T., 2010. Using ecological indices to measure economic and environmental performance of irrigated agriculture. Ecological Economics 69, 1731-1739.

Bellenger, M. J., Herlihy, A. T., 2009. An economic approach to environmental indices. Ecological Economics 68, 2216-2223.

Bellenger, M. J., Herlihy, A. T., 2010. Performance-based environmental index weights: are all metrics created equal? Ecological Economics 69, 1043-1050.

Bernhardt, E. S., Palmer, M. A., Alexander, G., Barnas, K., Brooks, S., Carr, J., 2005. Synthesizing U.S. river restoration efforts. Science 308, 636-637.

Bogetoft, P., Otto, L., 2011. Benchmarking with DEA and SFA. R package version 0.18.

Brisbane City Council, 2011. Norman Creek 2026 Vision and Concept Plan. Brisbane City Council, Brisbane. www.brisbane.qld.gov.au

Brisbane City Council, 2013. Norman Creek 2012-2031 Master Plan. Brisbane City Council, Brisbane. www.brisbane.qld.gov.au

Bunn, S. E., Abal, E. G., Smith, M. J., Choy, S. C., Fellows, C. S., Harch, B. D., 2010. Integration of science and monitoring of river ecosystem health to guide investments in catchment protection and rehabilitation. Freshwater Biology 55 (Suppl. 1), 223-240.

Charnes, A., Cooper, W. W., Rhodes, E., 1978. Measuring the efficiency of decision-making units. European Journal of Operational Research 2, 429-444.

Charnes, A., Cooper, W. W., Rhodes, E., 1981. Evaluating program and managerial efficiency: an application of data envelopment analysis to Program Follow Through. Management Science 27**,** 668-697.

Chessman, B., 2003. SIGNAL 2 – A scoring system for macro-invertebrate ('water bugs') in Australian rivers. Monitoring River Heath Initiative Technical Report no 31. Commonwealth of Australia, Canberra.

Cooper, W. W., Seiford, L. M., Tone, K., 2000. Data Envelopment Analysis. Kluwer Academic Publishers, Dordrecht.

Czyzak, P., Jaszkiewicz, A., 1998. Pareto simulated annealing – a metaheuristic technique for multiple-objective combinatorial optimization. Journal of Multi-Criteria Decision Analysis 6, 34-47.

Deprins, D., Simar, L., Tulkens, H., 1984. Measuring labour-efficiency in post offices. Chapter 10 in Marchand, M., Pestieau, P., Tulkens, H., eds., The Performance of Public Enterprises: Concepts and Measurement. North-Holland, Amsterdam.

Dyckhoff, H., Allen, K., 2001. Measuring ecological efficiency with data envelopment analysis (DEA). European Journal of Operational Research 132, 312-325.

Emerson, J. W., Hsu, A., Levy, M. A., de Sherbinin, A., Mara, V., Esty, D. C., Jaiteh, M., 2012. 2012 Environmental Performance Index and Pilot Trend Environmental Performance Index. Yale Center for Environmental Law and Policy, New Haven. www.epi.yale.edu

Engman, A., Ramírez, A., 2012. Fish assemblage structure in urban streams of Puerto Rico: the importance of reach- and catchment-scale abiotic factors. Hydrobiologia 693, 141-155.

Färe, R., Grosskopf, S., Lovell, C. A. K., Pasurka, C., 1989. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. Review of Economics and Statistics 71, 90-98.

Färe, R., Grosskopf, S., Lovell, C. A. K., Yaisawarng, S., 1993. Derivation of shadow prices for undesirable outputs: a distance function approach. Review of Economics and Statistics 75, 374-380.

Gironás, J., Niemann, J. D., Roesner, L. A., Rodriguez, F., Andrieu, H., 2010. Evaluation of methods for representing urban terrain in storm-water modeling. Journal of Hydrologic Engineering 15, 1-14.

Hajkowicz, S., 2006. Multi-attributed environmental index construction. Ecological Economics 57, 122-139.

Hajkowicz, S., 2007. Allocating scarce financial resources across regions for environmental management in Queensland, Australia. Ecological Economics 61, 208-216.

Hermoso, V., Pantus, F., Olley, J., Linke, S., Mugodo, J., Lea, P., 2012. Systematic planning for river rehabilitation: integrating multiple ecological and economic objectives in complex decisions. Freshwater Biology 57,1-9.

Hof, J., Flather, C., Baltic, T., King, R., 2004. Forest and rangeland ecosystem condition indicators: identifying national areas of opportunity using data envelopment analysis. Forest Science 50, 473-494.

Hughes, J. M., Huey, J. A., Schmidt, D. J., 2013. Is realised connectivity among populations of aquatic fauna predictable from potential connectivity? Freshwater Biology 58, 951-966.

Kuosmanen, T., Kortelainen, M., 2005. Measuring eco-efficiency of production with data envelopment analysis. Journal of Industrial Ecology 9, 59-72.

Lenat, D. R., 1988. Water quality assessment using a qualitative collection method for benthic macroinvertebrates. Journal of the North American Benthological Society 7, 222-233.

Llewellyn, D. W., Schaffer, G. P., Craing, N. J., Creasman, L., Pashley, D., Swan, M., 1995. A decision support system for prioritizing restoration sites on the Mississippi river alluvial plain. Conservation Biology 10, 1446-1455.

Lovell, C. A. K., 1995. Measuring the macroeconomic performance of the Taiwanese economy. International Journal of Production Economics 39, 165-178.

Lovell, C. A. K. Pastor, J. T., 1999. Radial DEA models without inputs or without outputs. European Journal of Operational Research 118, 46-51.

Macpherson, A. J., Principe, P. P., Smith, E. R., 2010. A directional distance function approach to regional environmental-economic assessments. Ecological Economics 69, 1918-1925.

Margules, C. R., Pressey, R. L., 2000. Systematic conservation planning. Nature 405, 243-253.

Meyer, J. L., Paul, M. J.,Taulbee, W. K., 2005. Stream ecosystem function in urbanising landscapes. Journal of the North American Benthological Society 24, 602-612.

Millington, H. K., 2013. Spatial analysis of the impacts of urbanisation on the health of ephemeral and perennial streams in South East Queensland. Griffith University, Nathan.

Palmer, M. A., Bernhardt, E. S., Allan, J. D., Lake, P.S., Alexander, G., Brooks, S., Carr, J., Clayton, S., Dahm, C. N., Shah, J. F., Galat, D. L., Loss, S. G., Goodwin, P., Hart, D. D., Hassett, B., Jenkinson, R., Kondolf, G. M., Lave, R., Meyer, J. L., O'Donnell, T. K., Pagano, L., Sudduth, E., 2005. Standards for ecologically successful river restoration. Journal of Applied Ecology 42, 208-217.

Paul, M. J., Meyer, J. L., 2001. Streams In the urban landscape. Annual Review of Ecology and Systematics 32, 333-365.

Peters, C. N., Marmorek, D. R., 2001. Application of decision analysis to evaluate recovery actions for threatened Snake River spring and summer Chinook salmon (Oncorhynchus tshawytscha). Canadian journal of fisheries and aquatic sciences 58, 2431-2446.

Ramírez, A., Engman, A., Rosas, K., Perez-Reyes, O., Martinó-Cardona, D., 2012. Urban impacts on tropical island streams: some key aspects influencing ecosystem response. Urban Ecosystems 15, 315-325.

Resh, V. H., Norris, R. H., Barbour, M. T., 1995. Design and implementation of rapid assessment approaches for water resource monitoring using benthic macroinvertebrates. Australian Journal of Ecology 20,108-121.

Rogge, N., 2012. Undesirable specialization in the construction of composite policy indicators: the Environmental Performance Index. Ecological Indicators 23, 143-154.

Roy, A., 2004. Can riparian forests mediate impacts of urbanization on stream fish assemblages?. The University of Georgia, Athens, Georgia, USA.

Sahoo, B. K., Luptacik, M., Mahlberg, B., 2011. Alternative measures of environmental technology structure in DEA: an application. European Journal of Operational Research 215, 750-762.

Scheel, H., 2001. Undesirable outputs in efficiency evaluations. European Journal of Operational Research 132, 400-410.

Sheldon, F., Peterson, E. E., Boone, E. L., Sippel, S., Bunn, S. E., Harch, B. D., 2012. Identifying the spatial scale of land use that most strongly influences overall river ecosystem health score. Ecological Applications 22, 2188-2203.

Steel, E. A., Fullerton, A., Caras, Y., Sheer, M. B., Olson, P., Jensen, D., 2008. A spatially explicit decision support system for watershed-scale management of salmon. Ecology and Society 13, 50-81.

Thompson, R., Parkinson, S., 2011. Assessing the local effects of riparian restoration on urban streams. New Zealand Journal of Marine and Freshwater Research 45, 625-636.

Tran, L. T., O'Neill, R. V., Smith, E. R., Wagner, P. F., Mehaffey, M., 2008. Watershed-based self- and peer-appraisal indices for integrated environmental assessment with a case study of the Mid-Atlantic region. Ecological Indicators 8, 308-315.

Walsh, C. J., Fletcher, T. D., Ladson, A. R., 2005a. Stream restoration in urban catchments through redesigning stormwater systems: looking to the catchment to save the stream. Journal of the North American Benthological Society 24, 690-705.

Walsh, C. J., Roy, A. H., Feminella, J. W., Cottingham, P. D., Groffman, P. M., Morgan, R. P. I., 2005b. The urban stream syndrome: current knowledge and the search for a cure. Journal of North American Benthological Society 24, 706-723.

Wenger, S. J., Roy, A. H., Jackson, C. R., Bernhardt, E. S. Carter, T. L., Filoso, S., Gibson, C. A., Hession, W. C., Kaushal, S. S., Martı´, E., Meyer, J. L., Palmer, M. A., Paul, M. J., Purcell, A. H., Ramı´rez, A., Rosemond, A. D., Schofield, K. A., Sudduth, E. B., Walsh, C. J., 2009. Twenty-six key research questions in urban stream ecology: an assessment of the state of the science. Journal of the North American Benthological Society 28, 1080-1098.

Zanella, A., Camanho, A. S., Dias, T. G., 2013. Benchmarking countries' environmental performance. Journal of the Operational Research Society 64, 426-38.

Zhou, P., Ang, B. W., Poh, K. L., 2006a. Comparing aggregating methods for constructing the composite environmental index: an objective measure. Ecological Economics 59, 305-311.

Zhou, P., Ang, B. W., Poh, K. L., 2006b. Slacks-based efficiency measures for modelling environmental performance. Ecological Economics 60, 111-118.

Zhou, P., Ang, B. W., Poh, K. L., 2007a. A mathematical programming approach to constructing composite indicators. Ecological Economics 62, 291-297.

Endnotes

[1] Cooper et al. (2000) provide a comprehensive survey of DEA and its uses.

[2] Macroeconomic applications are also popular, in part due to the availability of relevant OECD data. Zhou et al. (2006b) used a variant of DEA to study the environmental-economic performance of 30 OECD countries, and to calculate the opportunity costs of environmental regulations that constrain performance, using two inputs (primary energy supply and population), one desirable output (gross domestic product, GDP) and one undesirable output (CO_2 emissions). Zhou et al. (2007b) used a different variant of DEA to study the environmental-economic performance of 26 OECD countries, using two inputs (labour force and primary energy consumption), one desirable output (GDP) and four undesirable outputs (CO_2, SO_x, NO_x and CO emissions). Sahoo et al. (2011) used several specifications of DEA to study the environmental-economic performance of 22 OECD countries, using two inputs (capital and labour) to produce one desirable output (GDP) and one undesirable output (greenhouse gases).

[3] DEA provides a radial efficiency measure, and is not appropriate in the presence of undesirable outputs such as indicators of bad health. This leaves two options. We have chosen to convert indicators of bad health to indicators of good health prior to employing DEA to estimate efficiency. The alternative is to retain indicators of good and bad health and resort to a non-radial efficiency measurement method based on hyperbolic or directional distance functions. Färe et al. (1989) based their analysis on hyperbolic distance functions, while Bellenger and Herlihy (2009, 2010) and Macpherson et al. (2010) based their analysis on directional distance functions. Similar reasoning applies to the treatment of high stress and low stress indicators.

[4] Deprins et al. (1984) introduced the notion of dominance to production analysis, a contribution that was fleshed out in subsequent studies by Tulkens.

[5] An information server on composite indicators and ranking systems, not all of which use expert judgement to weight indicators, is http://ipsc.jrc.ec.europa.eu/?id=739.

[6] The first use of DEA to construct an aggregate index was Lovell (1995), who used DEA to aggregate macroeconomic indicators into an aggregate macroeconomic performance index for a number of Asian economies. Returning to the EPI, Zanella et al. (2013) developed alternative specifications of DEA to generate alternative aggregations of the 25 indicators in the EPI, and they compared their rankings with those in the EPI. Despite visibly different rankings, they found statistically significant positive rank correlations. Zhou et al. (2007a) are among many others who have used variants of DEA to aggregate environmental indicators to construct composite environmental indices across countries. At the macroeconomic level these environmental-economic output quantity indices are generalisations of Okun's Misery Index, Calmfors' Index and the OECD's Magic Diamond popularised by *The Economist*.

[7] Justification for this and other health indicators can be found in the Healthy Waterways 2006-2007 Annual Report at http://www.healthywaterways.org/EcosystemHealthMonitoringProgram/ProductsandPublications/AnnualTechnicalReports.aspx

[8] The software used to run the DEA models is available in the R package "Benchmarking with DEA and SFA," by Bogetoft and Otto (2011).