RICE UNIVERSITY

Sustainability Measurement: A New Evaluation Framework and a Case Study of Houston, Texas

by

Aaron Marc Robson

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Master of Science

APPROVED, THESIS COMMITTEE:

Pedro Alvarez, George R. Brown Professor, Chair, Civil and Environmental Engineering

Jim Blackburn, Professor of the Practice, Civil and Environmental Engineering

Phillip Bedient, Herman Brown Professor, Civil and Environmental Engineering

Stephen Klineberg, Professor, Co-Director, ~~Institute For Urban Research,~~ Sociology

HOUSTON, TEXAS
MAY 2011

ABSTRACT

Sustainability Measurement: A New Evaluation Framework and a Case Study of Houston, Texas

by

Aaron Marc Robson

Sustainability thinking has advanced considerably in the last few decades. Despite these advances, there remains a significant disconnect between thoughts and actions in this field. Sustainability metrics are one method proposed to help address this problem. A literature review of the most commonly used metrics and evaluation systems will first be presented, along with a critical evaluation of each system against the basic principles of sustainability. Next, a new evaluation framework will be proposed that addresses some of the shortcomings of the existing evaluation methods. This framework is designed to follow the best available thinking on a number of sustainability issues and overall, give a clear and concise evaluation of sustainability while still maintaining accuracy and scientific relevancy. Finally, a theoretical application of the framework to evaluating the sustainability of the city of Houston, Texas will be given, showing the necessary metrics required to follow the principles of the framework.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The concept of sustainability has enjoyed a rapid rise to prominence in recent decades. Since the release of the 1987 Brundtland Commission report "Our Common Future", which provided one of the first formal definitions of sustainable development in official policy, sustainability and all of its related aspects are becoming increasingly entwined in human society. This shouldn't be surprising – adhering to the basic principles of sustainable development will ensure a successful and prosperous future for human civilization on this planet. Very few people, regardless of their political, religious or social beliefs would argue against such a vision for our future.

Failure to properly consider these principles will likely have dire consequences, as history reminds us. Our past is littered with examples of great civilizations that came crashing down, in part, because they failed to understand the ecological limits of their immediate environment. The ancient inhabitants of the Fertile Crescent didn't understand the long-term effects of their irrigation methods and fell victim to waterlogged and salty soils (Jacobsen & Adams, 1958). The Mayans of Central America and the natives of Easter Island faced severe deforestation and top soil erosion (Brown, 1982 and Janssen & Scheffer, 2004) as a result of their agricultural and construction practices. Deforestation was also a main causal agent in the decline of the Norse civilizations of Scandinavia, Iceland and Greenland (Diamond, 2005).

Today, the technological progress and globalization of human society over the past few centuries means that a similar crash cannot be afforded – if our current civilization fails, it won't mean the ecological and social destruction of a few islands and a geographically isolated population. The civilization humans have built today touches every corner of the Earth. Every part of the globe will feel the impacts of the crash of human civilization.

The good news is that we have become increasingly aware of this uncomfortable possibility for our future and action is being taken to understand and hopefully avoid a devastating crash. The aforementioned Bruntland Commission report was born out of a formal effort by the United Nations to better understand the relationship between human society and the environment. What was developed was the notion of transitioning towards an indefinitely sustainable society through the process of sustainable development – codified in the formal definition: "sustainable development is development that meets the needs of people today without compromising the ability of future generations to meet their own needs" (WCED, 1987).

In 1992, the UN Conference on the Environment and Development (UNCED) in Rio de Janeiro (also known as the Rio Summit) took the rather nebulous ideas presented in the Bruntland report, and turned them into a set of definable policy goals, that become known as the Rio Principles. These outlined the changes in policy needed to meet the ideal of sustainable development. Though there had been some earlier efforts to address aspects of the sustainability problem, the Rio Principles were the first time the multi-dimensional nature of sustainability was truly acknowledged in official policy.

Since the Rio Conference, huge strides have been taken to better understand our civilization's interaction with the environment – how we exploit the natural resources provided by the Earth, how these resources are allocated and distributed throughout society, and how our societal structure addresses the innate needs of its inhabitants. Many different dimensions of sustainability have been identified though just three of them are universally referenced. These are: 1) The environmental dimension that relates to impacts of human activity on the natural environment; 2) The social dimension that deals with human needs – how the systems we have created are addressing various needs; 3) The economic dimension that is concerned with how resources are used to meet human needs. Various models have been developed to conceptualize how these dimensions affect sustainability including the commonly used three-legged stool, where each leg of the stool represents one of the dimensions and together they all support a sustainable future.

However, despite considerable advances in our understanding of the concepts behind sustainability, there remains a significant gap between scientific and academic ideals and political and economic reality. Put simply, we aren't doing a very good job of doing what we say we should be doing. This is due, in part, to the rather nebulous nature of sustainability. We can talk about the sustainability of everything from a piece of paper to the entire world. There is sustainable development, sustainability engineering, sustainable production, sustainable consumption and so on. What exactly it means to 'be sustainable' is unclear, and can depend greatly on the context in which the word is used. As such, it can be extremely challenging to ground sustainability principles in reality.

One way of addressing this problem is through the use of sustainability metrics and evaluation methods. The goal of sustainability metrics is to create effective, efficient and comprehensive methods of quantifying issues of sustainability. In doing so, we should be able to more easily make the critical link between important political, economic and social decisions and their effects on the sustainability of our civilization. This idea brings us to the purpose of this thesis.

It is the opinion of the author that the current metrics and evaluation methods for assessing and quantifying issues of sustainability are lacking, particularly those methods that attempt to assess sustainability in broad, holistic terms. This thesis attempts to address this issue. It consists to three major sections. In the first, current evaluation techniques are examined and evaluated in terms of how they adhere to some important concepts of sustainability. Then, based on the deficiencies identified in these methods, the second section presents a new framework for sustainability evaluation. Finally, to demonstrate how this framework can be used, the third section outlines the theoretical application of the framework to evaluating the sustainability of the city of Houston, Texas. This shows the kind of data needed to properly assess sustainability according to the principles of the developed framework. It is not expected that the method presented in this thesis will be perfect – however, the author hopes that it offers an incremental improvement on current methods, as well as perhaps a new perspective on the issues of sustainability that can be applied in the future to create even better methodologies.

PART 1: SUSTAINABILITY LITERATURE REVIEW

The intent of the first part of this thesis is to provide both an overview of the important concepts behind sustainability as well as a review of current sustainability measurements methods. These methods will be critically assessed on their ability to accurately reflect key issues relating to sustainability. These key issues will be the starting point for this section.

Concepts and Frameworks of Sustainability

The Three-legged Stool and Dimensionality

The three-legged stool was one of the first models to present a truly multi-dimensional picture of sustainability. Much of the initial thinking around sustainability centered around economic arguments, and later, environmental concerns. These centers of thought were not often merged, and social welfare concerns were rarely addressed directly. The idea of a three-legged stool model of sustainability was proposed to address these shortcomings. In this model, a sustainable society required consideration of each of the three supporting legs – economic, environmental and social. Neglecting any one of the legs meant missing key elements that are required to support a sustainable society.

Today, most environmental assessments make an effort to address multiple dimensions related to sustainability. The initial three dimensions have even been expanded to include institutional, cultural and other concerns (Shmelev & Rodriguez-

Labajos, 2009), though the distinctions between dimensional categories can be blurry. Regardless of the debate over minor semantics, there is a fairly uniform consensus that sustainability will require considering both the state of the natural and human systems, as well as the relationships and interactions between these two systems (Lozano & Huisingh, 2011). It is therefore critically important that any holistic sustainability evaluation method account for the multi-dimensional nature of sustainability by including metrics from each dimension. Multi-dimensional consideration is the first concept this thesis will assess evaluation methodologies against.

Carrying Capacity and Ecological Limits

One of the central tenets of sustainability is the idea of carrying capacity and ecological limits. This idea was what introduced sustainability thinking to the mainstream public with the environmental movement in the latter half of the twentieth century. Economist Herman Daly, with his concept of full-world thinking, and the Club of Rome in their 1972 publication "Limits to Growth" are some of the key developers of the concept of ecological limits, though countless other scholars have made important contributions in the field.

The thinking behind carrying capacity and ecological limits is that the Earth has a limited capacity to provide resources and absorb waste products and still maintain a relatively stable equilibrium. If resources are harvested at too high a rate, or too many harmful waste products are produced, the natural processes that provide for and clean up after human civilization will be overwhelmed. There are theoretical maximum values on how much humans can extract from the natural environment and how much waste can be expelled back without long-term negative effects – the sum of these values for the entire globe is considered the carrying capacity of Earth. It is possible to exceed the carrying capacity temporarily – to use a financial analogy, this means that we would be spending more than just the renewable interest on our money and eating into our capital reserve. Of course, this situation can only continue for so long until all of the initial capital is consumed.

In measuring sustainability, one can look at the relationship between human activities and the ecological carrying capacity. This relationship can be approached in two different ways. One can look at human impacts on the natural system, and decide what the maximum level of sustainable impact is. Alternatively, one can consider the products from the natural system that human civilization uses, and thus what the maximum level of use should be. An ideal system would likely uses elements of both viewpoints to create a more robust evaluation. At the very least, a good sustainability framework will recognize and address some concept of limits – this is the second criteria this thesis will use to assess sustainability evaluation methodologies.

Accuracy versus Simplicity

Devising an appropriate framework for sustainability involves a delicate balancing act between two opposing ideals. The concept covers extremely complex non-linear natural systems that are difficult to model correctly even for dedicated scientists using the most powerful supercomputers. An appropriate framework has to capture enough of this complexity to be sufficiently accurate. However, most of the policy-makers and politicians responsible for interpreting frameworks are not scientific experts. Sustainability covers such a wide range of disciplines anyways, that it is impossible to be an expert in all of them. This means that, at the very least, the end product of the framework has to be simplified – preferably enough to be understood by the average citizen.

How then can an appropriate balance of complexity be determined? Many indicator systems that use a single unit of measurement (discussed in detail later in this paper) have come under heavy criticism for misrepresenting key issues (van den Bergh & Verburggen, 1999; Pillarisetti, 2005; Dietz & Neumayer, 2007; Fiala, 2008). In order to reduce the very complex issues behind sustainability to a single dimension, many assumptions have to be made, some of which inevitably turn out to be incorrect, or misleading.

Conversely, systems that employ a lot of complex mathematics, or present an enormous number of final indicators are equally criticized. They end up being so complex that politicians are unlikely to trust the results, or they are simply unable to determine what results are relevant (Collins & Flynn, 2007). The solution to this problem that this paper proposed revolves around the arguments of weak and strong sustainability discussed in the next section. As will be discussed, these criteria are not set in stone, and to account for these changing ideas, a good evaluation system should also be flexible enough to permit tweaking of the balance between accuracy and simplicity in its measurements and reported data.

Weak versus Strong Sustainability

Weak and strong sustainability are two ways of thinking about sustainability from an overall systems perspective. Though they both claim to yield sustainable outcomes, they make very different assumptions to arrive at that conclusion. The concepts arose from various economic models constructed in the 1970s and 80s that were interested in resource depletion (Hartwick, 1977, for example). The models were based on the idea of natural capital stocks – the sum total of all components of value in the natural environment. This included both non-renewable and renewable resources, as well as ecosystem services, like waste assimilation. As the idea expanded, more types of capital stocks were added – manufactured, or economic capital, human capital (the value of the skills and knowledge of participants in an economy) and social capital (the value of the social relationships between people) (Kulig et al., 2010). These stocks became the basis for the sustainability framework known as the Hartwick-Solow model which, briefly, requires that total capital stocks are maintained over time.

As they relate to weak and strong sustainability, two schools of thought emerged about the maintenance of these stocks – in particular about the substitutability between different types of capital stock. One end of the spectrum argued for perfect, elastic substitutability between all forms of capital (Hamilton, 1994). This viewpoint became known as weak or soft sustainability. This approach is often adopted by economists, as it fits in well with the favored neoclassical economic theories. At the other end of the spectrum, substitutability between capital forms is completely forbidden – a perspective called strong or hard sustainability (Neumayer, 2003). This perspective is most likely to be adopted by conservationists and environmentalists, who are more concerned with human impacts on the natural environment.

An in-depth discussion of this debate is provided by Neumayer (Neumayer, 1999). This paper offers a brief summary of the critical points, particularly as they relate to sustainability evaluation. It is likely that there will be areas where some substitutability is appropriate, as well as areas where none is (Dietz & Neumayer, 2007). Past history has demonstrated that human beings are very adept at dealing with acute resource shortages and other similar limitations. Scarcity tends to drive innovation. This suggests that substitutability will be a valid assumption in related cases, such as resource utilization. There is also a case for cases of no substitution. Consider clean air – naturally occurring process in the environment continually purify the atmosphere. This is a service the natural system provides for no monetary cost to humans, and it is practically impossible for this service to be duplicated. Thus there is no form of capital that can completely substitute for clean air. This is the case for all of the 'life-support' functions of the environment (Barbier et al., 1994).

While there is no definitive consensus on where weak and strong sustainability should apply, strong arguments for certain categories do exist (Dietz & Neumayer, 2007). Raw materials used for production and consumption will have considerable substitutability. Waste assimilation capacity and natural amenity services can be substituted for to an extent. No realistic possibilities for substitution of the basic life-support functions provided by the environment exist, however. These are the guidelines regarding substitutability that this paper will follow.

In terms of the accuracy versus simplicity issue from the previous section, weak and strong sustainability can be used to decide when the system can be simplified without losing accuracy and relevance and when more complexity is necessary. The reasoning is that in cases where weak sustainability principles and substitutability are valid, complex systems can be condensed into a few key metrics and free-market forces used to determine the best solution within the constraints given by those metrics. In cases where strong sustainability must be used, individual indicators must be used for each separate dimension to maintain accuracy.

Without a definitive consensus on where weak and strong sustainability should apply, it is not a straightforward process to evaluate frameworks and evaluation methods on their use of these principles. In areas where there is a better defined consensus, a good evaluation system should follow these ideas. In more gray areas, the best solution is for the framework to pick a stance and make defensible arguments for its choice. As mentioned above, the ability to be flexible in how weak and strong sustainability apply will be instrumental to the long-term success of a framework. And ideal evaluation framework will follow these guidelines and establish an appropriate balance of weak and strong sustainability principles.

Framework Assessment Criteria

Four important concepts of sustainability have been presented. The basic ideas outlined in the preceding four sections will next be used to defensibly assess the merits of existing sustainability assessment methodologies. Consideration of multi-dimensional aspects, recognizing ecological limits, appropriate utilization of weak and strong sustainability and balancing complexity and simplicity in the methodology are the desired assets of an ideal sustainability evaluation system.

Single Unit Indicator Systems

These systems generally attempt to distill components of sustainability measurement into a single indicator with some form of 'universal' units. They are frequently billed as comprehensive – capturing a complete picture of sustainability in a single, easily comparable and relatable number. Various environmental, economic and societal issues are translated into figures with common dimensional units. These figures are then combined (usually through simple addition and subtraction) to yield a final 'sustainability score'. They differ from the aggregate indicators that will be discussed later in this thesis through the measurement units used in the scores. Aggregate indicators will generally use unit-less numbers, in contrast to single indicators that convert everything to a common unit. Below, the most common indicators of this type will be introduced and discuss some of the strengths and weaknesses of each method as they related to the evaluation criteria discussed previously.

Eco-efficiency

Efficiency has been a central tenet in modern economics since its inception. Eco-efficiency is simply the expansion of that idea further into the environmental domain. Efficiency in the traditional economic sense means utilizing available resources so as to maximize utility (measured by the production of goods and services). Applying this idea to the environmental sector, eco-efficiency can be defined as "creating more goods and services with ever less use of resources, waste and pollution" (WBCSD, 2000).

The concept was founded by the World Business Council on Sustainable Development in the 1990s (WBCSD, 1996; WBCSD, 2000) in their search for a term to describe sustainable development in business or economic terms. The intent of the term was to promote both economic and environmental benefits in industry by "maximiz[ing] resource productivity, gain[ing] bottom line benefits and reward[ing] shareholders, rather than simply minimiz[ing] wastes or pollution." (WBCSD, 1996) In this way eco-efficiency tries to address both the environmental and socioeconomic dimensions of sustainability.

Eco-efficiency has often been used as a means for evaluating sustainability in the academic community (Hoh et al., 2002; Montgomery & Sanches, 2002; Seppala et al., 2005; Mickwitz et al., 2006; Huppes & Ishikawa, 2009). Unlike some other index indicators, there is a wide variety of actual numeric indicators that fit in the category of eco-efficiency indicators. The common trait to all of them is that they attempt to compare resource usage (environmental input) against welfare provision (economic output) in a ratio type indicator.

Eco-efficiency indicators alone are unable to tell if the goal of sustainability has been achieved or not – a point that Hoh et al. acknowledges in their study of eco-efficiency indicators used in German Environmental Economic Accounting (Hoh et al., 2002). Mickwitz et al. also acknowledge that eco-efficiency is not a guarantee for a change towards sustainability (Mickwitz et al., 2006). Thus, the indicator does a poor job of interpreting ecological limits. Its greatest value is as a tool to be used within the context of other methods and frameworks that can account for limits and rely on eco-efficiency as a tool to focus improvement efforts.

The concept can theoretically cover multiple dimensions of sustainability, but how accurate the concept represents these dimensions can vary based on the exact metrics used. Both resources and welfare provision can be measured in many different ways, and some ways are better than others at addressing the dimensions of sustainability identified earlier. Eco-efficiency has the potential to address all dimensions of sustainability, but not with a single number – there is no one unit that can represent the economic, social and environmental dimensions of sustainability adequately.

The concept does have value in simplifying large numbers of complex measurements, but this application risks following weak sustainability principles in areas where strong sustainability ideas are more applicable, such as ecosystem services. Protecting against this contingency requires utilizing eco-efficiency within a framework that ensures that dimensions are kept separate when appropriate. To summarize, eco-efficiency ratios are useful tools, but their standalone value as evaluators of sustainability is limited.

Genuine Savings / Adjusted Net Savings

Genuine Savings, also referred to in the literature as Adjusted Net Savings, was a concept first presented by Pearce and Atkinson (Pearce & Atkinson, 1993) in response to a paper by Victor (Victor, 1991) that introduced the idea of natural capital as a measure of sustainability. Pearce and Atkinson expanded on that and devised an indicator to assess sustainability based on maintenance of natural and man-made capital stocks. Since these initial measurement efforts, the capital concept has been expanded to include human capital (the knowledge and skills represented in the population) and institutional or social capital (the value inherent in the social relationships in the population).

This initial indicator was defined by a sustainability index, Z, equal to the total savings minus the sum of the deprecations, or changes in natural and man-made capital in an economy (Pearce & Atkinson, 1993). Expanding this to include other forms of capital simply means adding in more depreciation terms, for whichever type of capital one desires to include.

By utilizing multiple types of capital stocks, Genuine Savings arguably covers a wide range of dimensions. How representative the measure is at accurately portraying each dimensions is highly dependent on how each form of capital is measured.

The primary value of the Genuine Savings indicator lies in its simplicity of presentation and understanding. By condensing all sustainability issues down into one 'savings' number, it is easy to compare different countries or regions. Tracking progress is theoretically very simple – if genuine savings is positive, the economy is considered sustainable. Moving towards a positive value would thus be symbolic of moving towards sustainability. In practice, things are not so simple. Choosing what component indicators are required to represent the different forms of capital is a difficult, subjective choice. Depending on the choice of indicators, data outliers and other anomalies can easily misrepresent the true picture of sustainability (Pillarisetti, 2005).

Since its inception, Genuine Savings has been extensively critiqued for its use of the weak sustainability argument. By simply summing all forms of capital, complete marginal elastic substitutability of capital forms is assumed – this shortcoming is fully disclosed by the original authors of the indicator (Pearce & Atkinson, 1993). The methodology has been modified over time to attempt to address this issue, but little progress has been made, largely because of the theoretical foundations of the method. Even new methods make critical assumptions that are unlikely to hold in practice (Muller, 2008). Certain forms of natural capital, in particular ecosystem services, are simply unlikely to be substitutable in the foreseeable future (Neumayer, 1999) which calls into question the validity of the Genuine Savings indicator.

Green Net National Production

Similar to the Genuine Savings indicator described above, green Net National Product (NNP) is based on the idea of compensating for the decline in natural capital stocks, following the Hartwick-Solow model of sustainability. NNP is obtained by deducting the explicit 'economic depreciation' of natural resource capital from the GDP of a country (Hartwick, 1990).

Many of the same strengths and weaknesses inherent to Genuine Savings can be applied to NNP. It does offer the clarity of a single representative 'sustainability' number. But the same weak sustainability arguments apply – natural capital is considered as a unified stock, and substitutable with growth in GDP. Increasingly, the scientific consensus is that this assumption is flawed (Dietz & Neumayer, 2007) thus limiting the validity of these type of indicators. However, this does not mean that weak sustainability indicators are irrelevant (Mota et al., 2010). Rather, weak sustainability can be used as a precursor to strong sustainability – subjects that fail to meet weak sustainability requirements will almost certainly fall short of strong sustainability.

An additional benefit to the NNP indicator is that it can be used as a measure of sustainable consumption (Onuma, 1999). This property arises because NNP is equivalent to a hypothetically maintainable 'constant consumption' level. Onuma proves this level is definable as the upper bound of constant consumption levels – an important property for evaluating economic policy, but as discussed above, lacking as a universal measure of sustainability.

Index of Sustainable Economic Welfare / Genuine Progress Indicator

This family of indicators began, like many economic-based indicators, as a method of adjusting GDP figures to account for environmental impacts. Lawn gives an excellent summary of how various accounting methods that deal with adjusted GDP have progressed (Lawn, 2007). The Index of Sustainable Economic Welfare (ISEW) indicator was born out of a paper by Nordhaus and Tobin (Nordhaus & Tobin, 1972) and later revisions and additions to that method by Daly and Cobb (Daly & Cobb, 1989). The idea behind the ISEW indicator was to comprehensively measure sustainable economic welfare. Here welfare means a complete measure of the utility of private consumers in the economy (Dietz & Neumayer, 2007).

The process takes numerical data on transactions deemed directly relevant to human well-being and adjusts these accounts further for various types of economic activity that traditional measures ignore (such as volunteer work and household chores) to create a complete list of influencing factors (Lawn, 2003). These factors can vary a bit from study to study based on data availability and individual preferences, but are relatively similar on the whole.

As with most economic-based indicators, there are a wealth of studies that have employed the ISEW and GPI as a measure of sustainability. This has typically been done at a country level, as data on most of the applicable factors is readily available from national governments (Wilson & Moffatt, 1994; Distaso, 2007; Bleys, 2008). There have been some attempts to apply the concept to smaller, local scales (Pulselli, 2006), but it is much more difficult to find adequate data.

The ISEW attempts to cover the full spectrum of dimensions relating to sustainability. Environmental degradation, resource depletion and in correlated fashion – carrying capacity – are accounted for in the calculations (though with questionable accuracy). In so far as it is intended as a measure of welfare, it covers social dimensions.

However, it relies on dubious neoclassical economic assumptions to make the link between its monetary-based metrics and 'real' measures of social welfare (Dietz & Neumayer, 2007). The method is probably most representative of the interactions between social and environmental systems as it examines material flows and how they related to a concept of welfare (not the most accurate one, but a concept nonetheless).

As with all the previous economic-based indicators, the ISEW completely supports a weak sustainability paradigm. By distilling environmental, social and interactive factors into a single unit of currency and using simple summation to come to an aggregate score, perfect substitutability is assumed in the model. This much is often admitted by study authors, who generally recommend supporting biophysical indicators to compliment the ISEW data (Neumayer, 1999; Lawn, 2007; Bleys, 2008; Brennan, 2008). Thus, as with the other economic-based indicators in this review, the ISEW is useful for its presentation of sustainability as a single number, but limited as a stand-alone indicator by its weak sustainability assumptions.

Ecological Footprint

The ecological footprint, introduced by Rees and Wackernagel (Rees & Wackernagel, 1994), was designed as an indicator to support strong sustainability principles. Their basic idea was to translate all human impacts on the natural system into productive land-area equivalents, and then compare the calculated required land area with the globally available land area. The footprint numbers can be meaningfully scaled down to sub-global regions by converting them to per capita figures and comparing that number to the global per capita footprint. This method creates a simple visual representation of the concept of ecological carrying capacity.

Substantial criticism has been levied at the methodology – much of which focuses on a perceived misrepresentation of the impacts of consumption (van den Bergh & Verburggen, 1999; Lenzen & Murray, 2001; Ferng, 2002; Fiala, 2008). The static nature of how ecological footprint calculations translate human impacts into equivalent land areas has been attacked as it fails to account for the effects of changes in production efficiency and emerging technologies (Ferng, 2002; Fiala, 2008). Equal physical weighting of all consumption impacts (a property implicit in the calculation methodology) doesn't necessarily correlate with the social weighting of those impacts, thus potentially leading to socially undesirable evaluation results (van den Bergh & Verburggen, 1999). There are also significant issues with comparing regional and global footprints that can lead to misleading conclusions about sustainability (Lenzen & Murray, 2001; Fiala, 2008).

The upside to all of this criticism has been the creation of a huge number of improved methodologies that fuse the ecological footprint concept with a wide range of complimentary environmental methods including input-output analysis (Wiedmann et al., 2006), local land use patterns (Ferng, 2001; Lenzen &Murray, 2001), scenario analysis (Ferng, 2002) and emergy analysis (Zhao et al., 2005; Siche et al., 2010b). The result of this has been a significantly more robust and useful indicator, albeit one that has increased greatly in complexity.

Despite these improvements, the ecological footprint remains of limited use as a stand-alone indicator of sustainability for two reasons. The first is that the scope of the ecological footprint is limited, and fails to cover all dimensions of sustainability. Even assuming that economic measurements are unnecessary for a snap shot evaluation of sustainability, the ecological footprint leaves out critical social aspects of sustainability. Take, for example, a country in which the per capita ecological footprint is well below the globally available value but does not adequately provide welfare for its inhabitants and has a socially damaging level of inequity. This would logically be an unsustainable situation, despite a 'positive' ecological footprint evaluation. Such a situation is, in fact, present in many developing nations.

The second issue deals with the familiar problem of substitutability and weak versus strong sustainability. The way the ecological footprint is structured, natural capital is completely separate from all other forms of capital meaning there is no substitutability between major forms of capital. This follows the more stringent strong sustainability principles as far as inter-dimension substitutability is concerned. The ecological footprint method does, however, allow for substitutability within the domain of natural capital, as it doesn't differentiate between different land area types in its aggregate calculation. As discussed earlier in this paper, this is a flawed methodology – some elements of natural capital are unsubstitutable even with other forms of natural capital and there is likely to be some substitutability between certain elements of natural capital and other capital forms (Neumayer, 1999; Dietz & Neumayer, 2007).

Environmental Space

Environmental space is a metric similar in nature to the ecological footprint in that it attempts to reconcile human resource usage with ecological limits. The indicator was introduced by the Wuppertal Institute in Germany, in collaboration with Friends of the Earth International in the mid 1990s (Mittler, 1999). Their idea was to focus on a set of key resources for which sustainable environmental capacity constraints could be identified and derive from these per capita resource usage targets. Each individual is assumed to be entitled to an equal share of the sustainably available resources.

A great deal less literature has been published on the subject of environmental space when compared to the ecological footprint. It is not immediately obvious as to why this is, but Mittler (Mittler, 1999) and Buhrs (Buhrs, 2004) identify some key shortcomings of the methods which might be to blame for its rather more modest adoption by the scientific community. The method is very dependent on the quality of data available – which can be lacking, even in highly developed Western nations (Mittler, 1999). Furthermore, the science that describes how the global ecosystem reacts to certain input and output levels is extremely complex and rife with uncertainty. Any limits defined by an environmental space analysis will be at most 'best guesses' (Buhrs, 2004).

Despite the limitations, there are some distinct advantages offered by the environmental space methodology. The actual calculation of resource usage is neither too complex, nor too simple – thus ensuring representative conclusions whilst still maintaining relative ease of use and calculation (Spangenberg, 2002). It is also helpful in setting concrete targets for various resource use reduction (Buhrs, 2009). A final and more unique benefit (though implicitly shared with the ecological footprint) is that it deals directly with the concept of environmental justice and equity (Mittler, 1999) by assuming that everyone is entitled to an equal share of the world's environmental space.

As discussed above, environmental space seems to strike a good balance between complexity and simplicity. It doesn't do as well at addressing all the dimensions of the sustainability relationship. As Spangenburg notes, environmental space does not directly address social or environmental issues related to sustainability (Spangenburg, 2002). These issues can be measured implicitly based on how the indicator is applied and what threshold values are set to – 'high' and 'low' scenarios that correspond to different levels of social value (Spangenburg, 2002).

Satisfying environmental space principles requires a high standard of strong sustainability. By evaluating each key input separately, the method permits no substitutability between components within its scope (largely environmental). This is likely too strong of a standard to be thoroughly realistic (Dietz & Neumayer, 2007). Still, the environmental space indicator offers some unique benefits, and strikes a better balance in many key areas than most single indicator systems.

Net Primary Productivity

This indicator is based around the concept of carrying capacity – it directly measures one aspect of ecological carrying capacity and treats that measurement as a proxy for other sustainability-related metrics. Net primary productivity (NPP) is defined as the the total amount of energy available to the planet's heterotrophs (consumers and decomposers) – measured by the total amount of solar energy that is fixed by plants minus primary producer respiration (Vitousek et al., 1986). As a sustainability indicator, the percentage of this energy humans are appropriating, as what is left over is the energy available to the rest of the species on the planet is of greatest importance. Leaving less and less energy for the rest of the planet will have ever increasing ecological impacts.

In 1986 it was estimated that close to 40% of potential terrestrial NPP was impacted by human activity (Vitousek et al., 1986). Since then, that figure has been revised and updated, as different measurement methods and more data have become available, but most estimates are in the 30-50% range (Haberl, 2006). NPP is not often used as a stand-alone indicator of sustainability, but typically as a metric within another indicator (e.g. Siche et al., 2010a; Peng et al., 2002). That said, how the concept relates to the principles outlined at the beginning of this paper can still be discussed – dimensionality, substitutability, simplicity versus accuracy, and ecological limits.

As NPP is intended to be a direct proxy of carrying capacity, it does a pretty good job of representing this concept, at least from a consumption perspective. It reduces all impacts down to how they utilize solar energy fixed by plants. The area where this method falls short is accounting for the waste assimilation part of carrying capacity. The waste products generated by human activities have huge impacts on natural systems and often overwhelm the capacity of the local environment to deal with them. This factor goes unaccounted for in NPP calculations except as energy flows.

NPP is openly a 'limited-dimension' indicator. It directly covers the environmental dimension, and indirectly measures elements of societal-nature interactions, insofar as the efficiency of meeting needs based on primary productivity appropriation is concerned.

The societal dimension is completely overlooked, and acknowledged as such (Haberl, 2006). This reason is why NPP is almost always used a merely a piece of a larger framework when evaluating overall sustainability.

By nature of its limited scope, NPP does not permit substitution between the primary forms of capital. It measures human impact on the natural system and thus, no amount of economic productivity, transferred knowledge or social relationships can offset declines in environmental health. The downside to the measure is that it aggregates all environmental impacts into one measure, when this may not be appropriate, as mentioned previously.

Aggregate Indicator Systems

Both aggregate and single indicator systems attempt to create, from their calculation procedure, a single sustainability 'score' that can then be used for comparisons or charting changes over time. The critical difference between the two methods is that aggregate systems use dimensionless final numbers, created by various mathematical techniques, while single indicator systems convert everything into a common unit which then allows easy aggregation. Dimensionless aggregation has its advantages and disadvantages, which will be discussed in detail for two major aggregation methods below.

Multi-Attribute Decision-Making

This is a relatively new method that takes advantage of fuzzy mathematics to perform aggregation on a suite of indicators. Fuzzy mathematics, related to fuzzy logic and fuzzy set theory, is a way of mathematically dealing with uncertainty and approximate values – in contrast with traditional logic (referred to as 'crisp' logic) and mathematics that require exact values. It has often been used in modeling 'real' systems, because most real-world features are not precisely defined, thus using the approximate variables of fuzzy math should give a more realistic picture.

Riberio (Riberio, 1996) gives a good summary of how fuzzy logic is used in decision-making tools. The decision-making process can be characterized as a process of choosing the best option amongst a set of alternatives to meet some specific goal. Uncertainty in the criteria that define that goal, and how each of the alternatives is evaluated against it, is carried over into the fuzzy model of the process. It is easy to see why sustainability modeling is a good application for this method – there is almost always uncertainty in the measurements and indicators used, and little clear consensus on what the end 'goal' of sustainability is.

One of the first models developed to assess sustainability using fuzzy logic was the Sustainability Assessment by Fuzzy Evaluation (SAFE) model (Phillis & Andriantiatsaholiniania, 2001). SAFE takes a suite of indicators, chosen based on data availability and a number of criteria regarding the attributes of the data, and normalizes each indicators according to statistical properties. These normalized values are then fed into a series of fuzzy rules that create groups of indexes, which are eventually retranslated by another set of fuzzy rules into an overall sustainability index. Kabak & Ulengin refine this model by using analytical tools to reduce redundancy in the indicators and simplify the fuzzy rule set (Kabak & Ulengin, 2007). There are many other studies available that use variations on this approach.

Fuzzy methods are highly dependent on the indicators that are initially chosen. The advantage they have is they are capable of using any indicator that has adequate data available for it – they don't require measurements be convertible to a standard unit. Thus for the purposes of assessing the fuzzy aggregation method, the best possible set of indicators is assumed to have been chosen which, by definition, would include indicators in all relevant dimensions of sustainability.

How well fuzzy aggregation addresses the issue of ecological limits is less clear. As per the earlier section of this paper, identification of a limit (hard or soft) that human activities are constrained by is required. Assuming that the set of indicators used in the calculation includes indicators that adequately measures the necessary areas, the only question is how well does the fuzzy concept convey the limits in these areas? This is dependent on the rules used in the aggregation process, but it is likely that the importance of remaining within limits would be diminished by the inclusion of so many other factors affecting the final index score. It is possible though, to weight the relevant indicators enough to keep proper semblance of meaning.

Fuzzy methods were designed to overcome the shortcomings of the single indicator systems in dealing with substitutability. In how the aggregation rules are defined, it is possible to reduce the impact of substitutability. The model used by Kabak and Ulengin actually permits substitutability when an indicator is above a certain threshold value, but not when it is below (Kabak & Ulengin, 2007). This is an interesting approach that is well suited to some parts of sustainability measurement, but in treating the intra- and inter-dimensional substitutability of all dimensions equally, is not as representative as a method that varies substitutability rules across the dimensions based on the current best consensus on what is appropriate.

With the end product of the fuzzy aggregation being a single number, interpretation and comparison of the final results are simple. However, the steps taken to arrive at this simple number at extremely complicated and different studies tend to use slightly different methods complicating things further. This is problematic for utilizing fuzzy aggregation in the real world – policy-makers and politicians need to be able to understand where numbers are coming from if they are going to trust them to base policy decisions on (Collins & Flynn, 2007). This is by far the biggest weakness of the fuzzy aggregation methodology.

Factor Analysis

Factor analysis is a mathematical technique that can be used to analyze interrelationships between large numbers of variables (Herrera et al., 2003). It is a method by which to condense information into a more easily understood package without losing the value of that information. To apply this method to sustainability evaluation, first a set of indicators is chosen. Then, after statistical normalization, the relationship analysis is performed. This yields a number of principal components that are dependent on the original set of indicators. Components with a sufficiently high statistical relevance are used as the final set of indicators.

This method can be used as a stand-alone indicator aggregation method, but this is not always the case. Factor analysis is often integrated into other procedures, such as fuzzy aggregation (Kabak & Ulengin, 2007). As an independent procedure, factor analysis has been applied to the local region of Baja California Sur in Mexico (Herrera et al., 2003). A general suite of indicators taken from the 1998 United Nations recommended list was broken down into principal components which were then summed to arrive at a final sustainability index. The method was also used to analyze a coastal area in China (Shi et al., 2003) – in this case, factor analysis was done to yield indexes for each identified dimension of sustainability (the traditional three of environmental, economic, and social). In a more limited scope, factor analysis was also used to condense biodiversity indicators in South Tyrol, Italy (Tasser et al., 2008).

As with the fuzzy aggregation method, factor analysis doesn't give us guidelines on which indicators to select for the suite. With this not being contingent on the actual methodology, a fully comprehensive set on individual indicators is assumed to have been selected. With no restrictions on what indicators can contribute to the index, factor analysis has the potential to fully represent the scope of sustainability.

The concept of ecological limits is again not directly addressed – there is no provision within factor analysis that puts limits on indicators relating to carrying capacity. These indicators are obviously included in the calculation procedure, and if they end up as one of the principal components of the index will have a large effect on the total score. But this is situationally dependent, not something inherent to the method. There are better ways of addressing ecological limits directly.

Unlike fuzzy aggregation, factor analysis does not strictly prohibit weak sustainability substitution. The method used to arrive at the set of principal components does involve summing the contributions of all the indicators, though it is not entirely clear that this violates strong sustainability, as this process doesn't result in a final sustainability score. After principal components have been decided, how they are used is up to the individual evaluator. In most cases though, principle components are simply summed, leading to substitutability issues. Overall, factor analysis is less representative of strong sustainability than fuzzy aggregation.

The calculation procedure that factor analysis uses to condense the original set of indicators down into its principal components is not simple – it uses a fair bit of matrix algebra. It is, however, simpler than the fuzzy logic rules uses in fuzzy aggregation. Furthermore, the procedure maintains more of the original indicators used, and is thus more transparent in regards to where the final scores are coming from. This makes it more likely to be trusted by politicians and the like, even if they do not understand the mathematical foundations on which the method is based.

Multiple Indicator Systems

The final type of evaluation framework that this paper will consider is multiple indicator systems. These frameworks are characterized by a set of individual indicators, which may range from very specific (atmospheric ozone concentrations, for example) to very broad, even possibly utilizing some of the single indicator systems described previously. These frameworks also do not use any sort of aggregation procedure. The results of each indicator are presented separately, and it is left up to the viewer to make an overall sustainability assessment, if required.

Multiple indicator systems are far and away the most popular framework for sustainability evaluation. There are hundreds of published assessments based on multiple indicators in the academic and general literature. Most major cities across the world have developed some form of indicator-based assessment and there are a large number of country-level assessments available as well. With such a breadth of data, a complete analysis of each evaluation would be well beyond the scope of this paper. Instead this paper will focus on evaluating the general methodology of multiple indicator systems.

With such variability in the types of multiple indicator systems currently used, how accurate and representative a framework is will be highly dependent on the indicators selected. One of the downsides to multiple indicator systems is that unlike many of the single indicators, there is no guidance provided on what should be measured, which can lead to lopsided and inaccurate evaluations if the evaluators do not properly consider the important principles of sustainability. If it is assumed that an appropriate suite of indicators has been selected, then a multiple indicator framework will allow for an excellent representation of all the relevant dimensions of sustainability.

The flexible nature of using multiple indicators also means that it can be very easy to accurately depict ecological limits and carrying capacity. One simply has to select one or more indicators representative of local carrying capacity and when evaluating this indicator define appropriate limits that cannot be exceeded without reflecting an unsustainable overall system state.

Again, the flexibility offered by multiple indicator systems means that the framework can be very finely turned to get the optimal balance of weak and strong sustainability. Dimensions where weak sustainability substitutability is appropriate can be covered by one of the many aggregate indicators available, offering the simplest perspective possible. For dimensions that require strong sustainability principles, more specific indicators can be used to preserve substitutability concerns. The result is a framework that is as simple as possible while still recognizing weak and strong sustainability arguments.

Finally how well multiple indicator systems balance complexity and simplicity is considered. Many of the frameworks used currently fall into the trap of attempting to measure too many things. Some systems have upwards of 100 indicators and any meaningful information gets drowned in a sea of data. However, if the guidelines of weak and strong sustainability outlined above are followed, the multiple indicator system can be optimized to present conclusions in the most efficient manner. The biggest downside to multiple indicator systems is that comparisons between different areas and over time for the same area are very difficult. If over ten years, three indicators have gotten better and four have gotten worse has the overall sustainability increases or decreased? It is very difficult to say. But is that question really important? Arguably, anything that isn't completely sustainable is undesirable and thus still requires improvement. However, feedback on the impact of policy decisions is critically important to move toward sustainability – feedback that can be difficult to extract from multiple indicator systems.

PART 2: A NEW SUSTAINABILITY EVALUATION FRAMEWORK

The preceding section outlined a number of the flaws of current sustainability evaluation methods. Though many of the methods offer unique advantages, and can be extremely useful for an assortment of specific applications, most tend to fall short as general sustainability evaluation tools. The following section of this paper will propose a new evaluation framework designed to address these shortcomings and provide a comprehensive measure of sustainability.

The framework will first be presented in a general format, covering everything from the broadest questions of sustainability down to a set of measurement categories that will be applicable to any type of sustainability evaluation. Then, to demonstrate how the framework can be used to assess sustainability of a target system, a hypothetical evaluation of the sustainability of a major US city (Houston, Texas) will be presented. In this procedure, a suite of indicators that will comprehensively measure each of the categories from the general framework will be defined. The numerical data for the suite of indicators will not be included – in fact, it is likely that data will not be readily available for some of the chosen indicators. Rather, the hypothetical process will reveal the kind of things that should be measured to properly assess sustainability.

Since actions towards sustainability are often best implemented at the local level (citation), and local regions will differ greatly in the exact metrics that are both relevant and have meaningful data available, the final suite of indicators for Houston should not be copied without thought into other sustainability assessments. Only the general framework presented is universally applicable. How best to represent each of the categories from the framework is a question that must be answered by each individual evaluator. The evaluation of Houston is intended to simply demonstrate the logical process needed to customize the general framework to a specific target.

A Rational Basis for the Framework

The presentation of the sustainability evaluation method begins by laying out a rational basis for the structure of the framework and considering the problem of sustainability at its most fundamental levels. Many studies that have looked at this problem have described sustainability as an interdependent relationship between human society and the natural environment (Vucetich & Nelson, 2010). At this level then, sustainability is a very simple problem – if the state and trends of the natural environment and the human society are sustainable, then humans are living 'sustainably'.

Of course, this assertion is not particularly helpful without some idea of what 'sustainable' means. Ultimately, the natural environment will continue to exist, regardless of human impacts, until the sun ceases to function. Thus, talking about how 'sustainable' it is can be a bit confusing unless a context for it 'sustainability' is defined. The context that will be used in this paper is that of a life-support system for human existence – thus sustainability of the natural system requires it to be able to support human life.

The 'sustainability' of human society should be similarly qualified. Humans have lived on this planet for millions of years as hunter-gatherers and if our society voluntarily chose to return to this lifestyle humans could probably continue living for the indefinite future. Barring cataclysmic disaster however, this is not a feasible course of action – the majority of modern society is not going to willingly give up the comforts and conveniences of modern life. Thus, the sustainability of the 'modern' society must be considered, which can be conceived as a society being able to provide non-declining welfare to its inhabitants from the present day, measured in terms of happiness, life satisfaction, life expectancy, infant mortality and other associated measures.

With an understanding of what sustainability means for these two initial dimensions, a more detailed framework can now be gradually built up. It must thoroughly cover the first two dimensions of the sustainability problem and determine if the necessary conditions for natural system and human system sustainability are being met. Based on this assessment, a yes-no type evaluation of the target's sustainability can be produced. However, sustainability is about more than simply if a system is sustainable or not. Some insight into the mechanisms that drive a system towards or away from sustainability is desirable. For this, the relationship between the natural and human systems must be examined. If the method by which human impacts on the natural environment are translated into societal well-being can be described then reasons why certain elements of the natural and human systems appear unsustainable in the initial analysis will be uncovered.

The final structure of the evaluation framework is as follows: three separate dimensions will be evaluated to determine an overall assessment of sustainability. The natural system and the societal system are evaluated to determine the extent of 'sustainability'. Then analyzing the interactions between the two systems will reveal the mechanisms behind the target's sustainability or unsustainability and lead to recommendations for appropriate courses of action that will move the system towards sustainability. A visual representation of the structure of the framework is provided in Figure 1, below.

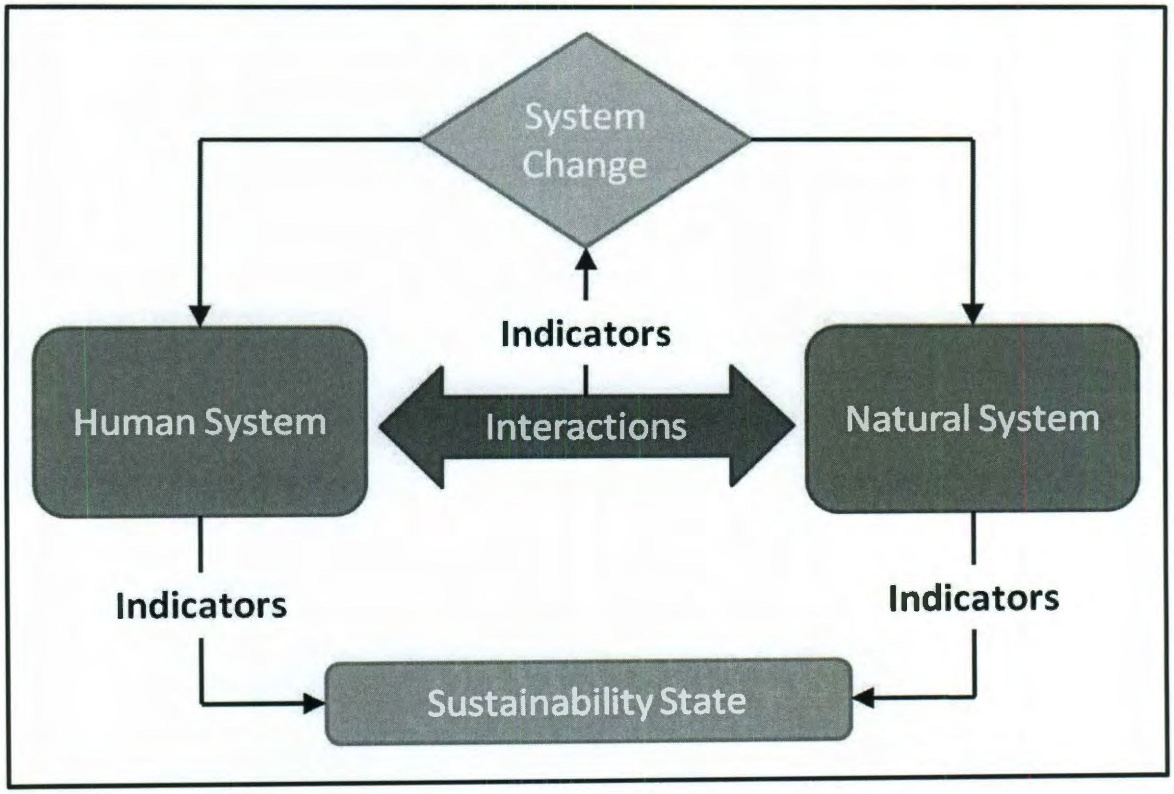

Figure 1: Framework diagram depicting relationships between categories

Dimensions of the Framework

The preceding section described the logic behind the framework, and introduced the three primary dimensions that will be used to evaluate overall sustainability. Next, these dimensions will be broken down into separate sub-dimensions that can be directly measured though various indicators and metrics. This process will be guided by the definitions provided previously on what constitutes a sustainable system.

The Natural System

The sustainability of the natural system is defined by its ability to support human existence. Thus, the sub-dimensions of the natural system will be based on the different ways that humans use the natural environment to support themselves. Foremost, and arguably most importantly, the natural system provides basic life-support functions, which are called ecosystem services. Exactly what these are can be difficult to define – Gretchen Daily offers a simple and easy to understand conceptualization that this paper adopts. If a colony were established on the moon (an environment that provides no life-support functions) clean air, water and a host of other things that the Earth naturally provides would have to be artificially supplied – these are ecosystem services (Daily, 1997). As humans are completely dependent on these services for survival, our actions must not adversely impact the provision of these services. A sustainable community will maintain these services at adequate levels.

The next sub-dimension is physical resources. This covers all of the resources extracted (both renewable and non-renewable) for human consumption. Humans utilize these resources for almost everything we do as a society – from building cars and houses, to eating, to transportation. Renewable resources such as timber and fish must not be overused. Sustainability requires that the natural system be capable of continuing to provide these materials that contribute to human welfare, or in the case of non-renewable resources allow time to develop appropriate substitutes so that welfare is not affected. If non-renewable resources become prohibitively expensive to extract without a viable substitute available, there will be some decrease in human welfare associated with that loss, which violates the previously stated sustainability definition.

The final sub-dimension of the natural system is biodiversity and natural ecosystem health. This dimension does not have as direct an effect on human welfare as the preceding two dimensions, and its inclusion in this framework is likely to be more heavily criticized. There are good reasons for its inclusion, however. The resilience of the natural environment to external shocks, such as sun activity, volcanic eruptions and climatic shifts is tied strongly to biodiversity. More biodiversity in an area means that it is more likely to survive such events with its critical ecosystems intact – ecosystems that humans are likely to depend upon in some fashion. Furthermore, the simple aesthetic beauty of natural environments has a large positive effect on the mental well-being of people. For these reasons, a sustainable community will maintain biodiversity and ecosystem health in its area.

The Human System

The primary concern of modern society, as it relates to the previous definition of societal sustainability, is the provision of welfare. Welfare provision is the first sub-dimension of the human system. A sustainable society must be able to provide at least the same level of welfare to its future inhabitants as it does its current ones. A system that sacrifices the needs and welfare of future generations to serve the present will ultimately die off – the countless skeletal remains of failed civilizations scattered across human history attest to this fact. In evaluating this dimension, careful consideration should be given to the definition of welfare. Economic systems use money, or income, as a proxy for human welfare. This concept, on the whole, is flawed. In evaluating welfare provision in this framework it is important to obtain more direct measures of welfare.

In addition to providing essential welfare, there is strong consensus that a sustainable society must consider social justice, or equity. Despite this consensus, the reason for including equity in sustainability evaluation isn't immediately obvious. As long as all members of a society have an adequate level of welfare provided to them, nothing else should be required. However, there is strong evidence to support the notion that humans tend to view themselves on a relative scale to their peers, rather than an absolute scale (citation). Someone who is drastically 'poorer' by some measure than most of his or her peers is unlikely to be satisfied with their situation, and likely to attempt to remedy it. Thus, a society with extreme differences in welfare is likely to harbor strong feelings of discontent in parts of its population and will not be sustainable in the long run until this inequality is diminished.

System Interactions

The dimension of system interactions deals with how the human system manages inputs from the natural system (i.e. natural resources) and outputs to it (i.e. waste products) to perform its vital functions of welfare provision and equity. It is in these interactions that the answers to why various parts of the preceding two systems appear sustainable or unsustainable will be found. Three categories of factors that govern this relationship can be identified.

Economic factors are the first sub-dimension of system interactions. These factors deal with how society deals with the production and distribution of goods and services. This covers how various natural resources a community extracts are utilized, who and where these resources go to, and what becomes of the products of these resources. All questions relating to why some component of welfare is not being met, or why some element of the natural system is being degraded will ultimately run through these economic factors.

Technological factors are very similar to economic factors – this paper make the distinction that these factors are solely concerned with how to improve the efficiency of the various methods that have been decided upon through economic factors. Of course, improvements in technological efficiency will feedback into economic-based decisions on what processes to utilize for various purposes, but the efficiency of the chosen processes can still be tracked to determine where improvements can be made.

Finally, the third sub-dimension concerning system relationships is one of outcome considerations. There is likely to be a wide range of sustainable end-state scenarios for human societies that range from harshly exploitive of the natural environment to richly harmonious with it. There isn't necessarily one 'right' end-state that all communities should strive to achieve – as long as a particular end-state is sustainable, it has just as much intrinsic worth as another. Each community must come to a collective decision about where on the spectrum of scenarios it wants to land – the fundamental economic factors that make up the first sub-dimension of system interactions depend greatly on this decision. Furthermore, it is important that this discourse be an ongoing process, as people's attitudes and ideas will change based on the effects of decisions previous made (Vucetich & Nelson, 2010). Consequently, the sub-dimension of outcome considerations will measure the degree of participation in the community for collective decision-making.

Overall Framework

This paper has so far described a general framework of three major dimensions – the natural system, the human system, and system interactions – that can be used to comprehensively evaluate sustainability. Figure 2, below highlights each of the sub-categories within the framework.

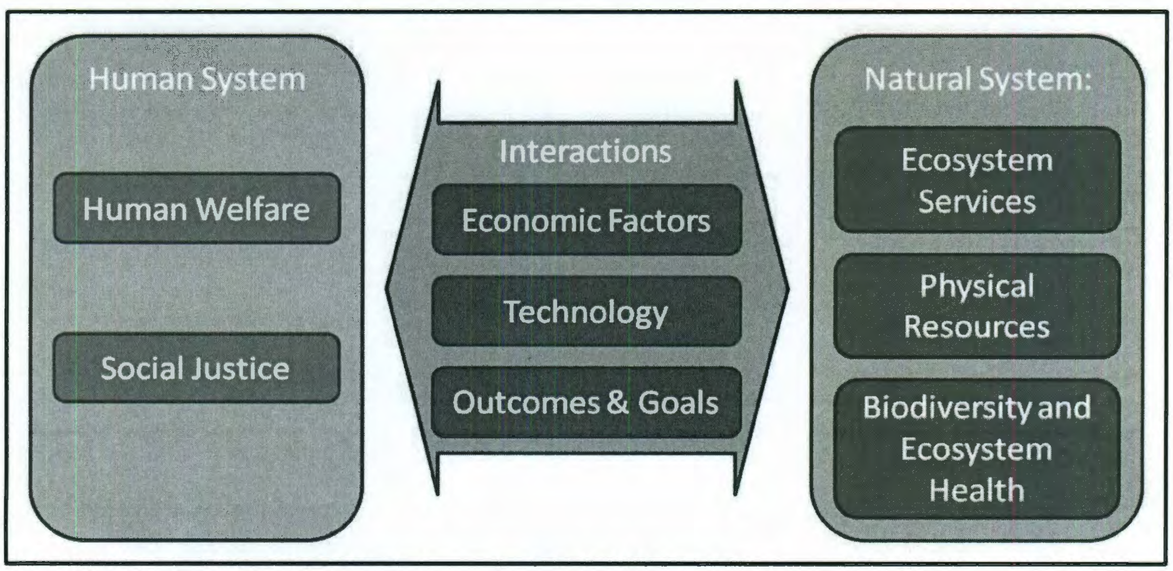

Figure 2: Framework categories

The natural system consists of sub-dimensions of ecosystem services, natural resources, and biodiversity and ecosystem health. The human system requires sub-dimensions of welfare provision and social equity. Economic, technological and outcome considerations affect the dimension of system interactions. Each of these sub-dimensions requires evaluation to determine overall sustainability status and future direction. If any one sub-dimension of the natural and human systems is identified as unsustainable, the entire system is unsustainable – these dimensions act as separate components, adhering to strong sustainability principles insofar as substitutability between different dimensions is not permitted. No amount of welfare can offset a failure in the provision of vital ecosystem services, for example. If any particular sub-dimension fails the assessment the system interaction variables must then be looked at to determine the cause of this problem.

This framework can be applied to any area-based sustainability evaluation. To illustrate this process, the following section of this paper will apply the principles of the general framework to the specific needs of the city of Houston, Texas. In doing so, a specific set of indicators will be created that the author thought was relevant to the target area to represent the dimensions of the general framework. How these indicators were determined can serve as a guide to future evaluators wishing to utilize this framework. It will also give an idea of the types of data required to properly assess sustainability, not all of which will be available.

PART 3: INDICATORS FOR HOUSTON

The Natural System

Ecosystem services constitute the first sub-category within the natural system category. There are seven identifiable major services: air purification, water purification, flood and drought mitigation, waste decomposition, soil generation and renewal, plant pollination and finally, climate stabilization (Daily, 1997). Each of these services is outside the technological ability of humans to provide artificially on the scale that the natural environment does, and for a comparable cost (free). Thus, the measurement methodology considers all ecosystem services as subject to the non-substitutable constraints of strong sustainability principles, up to the threshold required for sustaining human life. This requires that the provision of the basic levels of these services be independently measured and reported. This paper did not include measurements of all of these services in the assessment of Houston, as not all are present within the geographical boundaries of the region. Soil fertility within the city is a relatively minor concern – very little productive agriculture occurs in the city. Linked to this is plant pollination, which is also not of direct importance within the city Houston. Thus, the final list of services of concern to Houston is: clean air, clean water, flood mitigation, waste decomposition and climate stabilization. For each of these services, an acceptable threshold for sustaining human life must be identified, and the region assessed on whether or not its activities are impacting the ability of the environment to provide that level of services.

The second sub-dimension is physical resources. Historically, physical resources have proven to be largely substitutable – the economic system is structured such that when a resource becomes scarce, monetary incentives push for the development of alternative technologies. By allowing for substitutability, an aggregate indicator can be used for this sub-dimension, which simplifies the overall evaluation. The nature of cities (a high population density area supported by low density, productive “hinterlands”) means that directly comparing Houston's physical resource usage to the resources provided by the physical land area occupied by the city is meaningless. Instead, per capita physical resource use is considered, and compared to globally available amounts. Houston will only be sustainable if it does not exceed its share of globally available resources. The aggregation procedure by which both Houston's usage and globally available amounts are counted must be consistent to ensure that comparisons are meaningful. An indicator similar to the ecological footprint, environmental space, or material-flow accounts would be appropriate for this assessment.

The final sub-dimension of the natural system is biodiversity and ecosystem health. Despite being a heavily populated city, Houston still contains large tracts of natural habitat home to a wide variety of species. For this assessment, the degree of overall genetic biodiversity is of greatest concern, indicators of which can be found in various biological studies. Habitat area is also important for both maintaining diversity, and for its aesthetic value. An area-based indicator could be used to evaluate this point.

The Human System

Relevant indicators in the sub-dimension of human welfare will vary greatly between different locales, because different areas face very different challenges in this sector. A city in a developing nation in Africa would likely be very concerned with access to clean, fresh water, nutritious food and basic sanitation services. In Houston, the overwhelming majority of the population already has access to these services, so measurement efforts can focus on other services that weren't as widespread in the city. This helps to simplify the overall evaluation results, by eliminating superfluous data. To cover provision of basic welfare, a basic basket of goods, comprised of all the expenditures necessary to live in Houston is used. This includes at least food, shelter, clothing, health-care, and transportation. The cost of this basket is then compared to income levels in the city – anyone who does not make at least this level will find it very hard to meet this basic welfare needs. There are two other quantitative categories that should be covered in Houston – educational performance and health-care. In addition to the preceding quantitative welfare measures, assessment of general human welfare should include qualitative measures that look at things like life satisfaction and happiness. These measures are likely to come from city-wide surveying efforts.

The simplest indicator of social equity is income inequality, or the income gap between the richest and poorest members of society. Even though increased income is not a guarantee of increases happiness or satisfaction, it does mean that one will have better access to the vast majority of goods and services. Extreme inequality in the degree of access to these things will inevitably breed discontent amongst the disenfranchised. In assessing social equity, reducing the income gap towards a sustainable threshold is a primary concern. Secondary factors in this dimension are things that relate to specific concerns within the city of Houston. There is a huge disparity in the quality of education available to people living in different areas of the city, which a sustainability evaluation should consider. Crime rates also vary greatly by area – the goal of low overall crime rate, should be expanded to include low variability between different parts of the city. Finally, there are many areas in Houston where inhabitants are subject to disproportionate impacts from various environmental issues – a concept called environmental justice. A sustainable Houston will not only reduce total environmental impacts and problems, but ensure that there are not areas facing impacts that have been lost in the aggregate evaluation.

Systems Interactions

The dimension of system interactions is intended to be used as a tool to give insight into the sustainability evaluation results from the first two dimensions. Without actual data to draw results from regarding the state of the natural and human systems, it is difficult to decide exactly what parts of the interactions should be focused on. Having said that, there are some general guidelines that will be useful in choosing focus areas after assessment data has been reviewed.

Many economic factors are already implicit in many of the other dimensions already identified in the framework, which is why it plays such a central role in determining sustainability. If ecosystem services is identified as a deficient category, it might be because economic models do not consider their value when determining the 'optimal' situation. Similarly, if certain welfare measures are failing, despite increasing income levels, a reasonable conclusion might be that using income to measure utility in economic models is a flawed idea. This is the kind of thinking that allows levers of change for various environmental and social issues to be identified.

Technological-based measures allow areas of inefficiency to be identified. Essentially, technology allows communities to decouple from the relationship between increasing welfare and increasing resource consumption and waste generation. It allows more welfare to be gleaned from less consumption. Different sub-dimensions of the framework can be analyzed based on how well a particular resource is being used to provide welfare compared to others. Areas that require high utilization of the natural system to provide relatively less beneficial welfare are areas to target for improvement to move towards sustainability.

Finally, with outcome considerations, the actual end goals for various indicator categories can be determined. It is important to ensure that the relevant discourse is actually taking place – in a democratic society, much of this dialog will occur in various political institutions. Some measure of this should be taken. Furthermore, it would be extremely helpful to get some sense of where the general populace thinks Houston should score in the various indicator dimensions – this will help guide us towards sustainable solutions to the various issues that are agreeable to all.

Final Indicators

With all of the considerations for how to choose specific indicators laid out in the preceding sections, the final draft of indicators for Houston is presented on the following page in Table 1. These might not be the ideal set of indicators for Houston, but are at least a relatively realistic set that could be measured. Continuing with that point, a second table (Table 2) is presented that analyzes the quality of the data available for each of the chosen indicators. As is shown, there is great variability in the quality of data available for the selected indicators. Much of the data is rated Fair or Poor, which means that it cannot be tracked or analyzed in its current format (either because it isn't measured, or it requires a great deal of processing to present it in a digestible format). This serves to highlight how much more effort needs to be expended on collecting, tracking and analyzing data related to sustainability performance.

Macro-category	Sub-category	Component	Indicator
Environmental	Eco-System Services	Air Purification	Person-days of exceedance of PM standard
			Person-days of exceedance of ozone standard
		Water Purification	Ideal freshwater inflow ratios
			Turbidity or bacteria levels in area streams
		Flood Mitigation	% of area inundated by 100-year design flood
		Waste Decomposition	Dissolved oxygen concentrations in receiving streams
		Climate Stabilization	Tonnes of CO2 equivalent emissions per capita
	Biodiversity	Species Distribution	Change in diversity and size of wildlife populations
		Habitat Area	% various land types impacted by development
	Physical Resources	Aggregate Usage	Material flow accounts for renewable and non-renewable resources
Societal	Welfare	Purchasing Power	Cost of standard basket of goods versus median income
		Healthcare Access	Infant mortality
			Healthcare spending per capita
		Education Access	High school graduation rates by area
		Life Satisfaction	% of population responding that they are 'very satisfied' with their life
	Social Justice	Inequality	Difference between 25th and 75th income percentiles
			% of population over 25 with advanced degrees
		Crime	Various crime rates by area
		Environmental Justice	% of population living within 1/2 miles of toxic sources
Interactions	Economic	GPI	GPI calculation
		Diversity	Divesification of GPI across different sectors
	Technological	Eco-efficiency	EF versus GPI
		Green Investment	$ investment in various green technologies
	Outcomes	Voter Participation	Voter turnout numbers
		Community Involvement	% of population involved in altruistic activites

Table 1: Selected indicators for Houston

Indicator	State of Data	Rating
Person-days of exceedance of PM standard	Base data available, minor processing needed	Good
Person-days of exceedance of ozone standard	Base data available, minor processing needed	Good
Ideal freshwater inflow ratios	Full data set available	Good
Turbidity or bacteria levels in area streams	Base data exists, needs to be acquired and processed	Fair
% of area inundated by 100-year design flood	Base data available, minor processing needed	Good
Dissolved oxygen concentrations in receiving streams	Base data exists, needs to be acquired and processed	Fair
Tonnes of CO2 equivalent emissions per capita	Full data set available	Good
Change in diversity and size of wildlife populations	Significant new measurement efforts required	Poor
% various land types impacted by development	Base data exists, needs to be acquired and processed	Fair
Material flow accounts for renewable and non-renewable resources	Significant new measurement efforts required with extensive data processing	Very Poor
Cost of standard basket of goods versus median income	Base data exists, needs to be acquired and significant analysis completed	Fair
Infant mortality	Full data set available	Good
Healthcare spending per capita	Base data exists, needs to be acquired and processed	Fair
High school graduation rates by area	Full data set available	Good
% of population responding that they are 'very satisfied' with their life	Preliminary data is available, but a larger and more detailed set is needed.	Fair
Difference between 25th and 75th income percentiles	Full data set available	Good
% of population over 25 with advanced degrees	Base data exists, needs to be acquired and processed	Fair
Various crime rates by area	Full data set available	Good
% of population living within 1/2 miles of toxic sources	Full data set available	Good
GPI calculation	Base data exists, needs to be acquired and processed	Fair
Divesification of GPI across different sectors	Base data exists, needs to be acquired and significant analysis completed	Poor
EF versus GPI	Significant new measurement efforts required with extensive data processing	Very Poor
$ investment in various green technologies	Base data exists, needs to be acquired and processed	Fair
Voter turnout numbers	Full data set available	Good
% of population involved in altruistic activites	Significant new measurement efforts required	Poor

Table 2: Quality of indicator data

CONCLUSION

This thesis has attempted to first, demonstrate the need for a new approach to sustainability evaluation by outlining some of the major flaws of current methods, and second, propose a new approach that can be built upon in the future. When evaluated against some of the basic principles of sustainability, it is obvious that the single indicator systems commonly used in academic and political settings fall well short of ideal. Multiple indicator systems, both aggregate and separate offer greater potential, but require a great deal of thought on exactly what they are going to measure.

The new framework this thesis proposes if far from the finished article. It will, however, hopefully serve as an excellent starting point for the creation of more robust and detailed frameworks. It highlights the importance of careful consideration of the weak versus strong sustainability debate, balancing accuracy and simplicity, recognizing ecological limits and understanding the complex dimensionality of sustainability. Its application to measuring the sustainability of Houston demonstrated the magnitude of the challenge of appropriate sustainability evaluation. Many of the desired metrics that would appropriately fit the framework are simply unavailable at the present time, or data on them is woefully incomplete.

With these points in mind, this thesis would like to make some recommendations for continued work in this area. This framework was developed in a purely academic setting – the most important step in moving forward would be to get constructive input from political and economic decision-makers, and revise the framework with their help. Following this, it would be possible to make a real, practical use for the framework in a local setting. If local leaders are committed to affecting change in how their region approaches sustainability, this framework can move from simply being an academic exercise to a practical tool. This will require the development of new metrics that appropriately relay the desired information, and the creation of tools that relate real-world policy decisions to their effects on the various areas of the framework. Only such a collaborative effort, in which all parties are committed to changing how things work, will help move society towards a more sustainable future.

REFERENCES

1. Barbier, E. B., Pearce, D. W., Markandya, A., 1994. Environmental sustainability and cost–benefit analysis. Environment and Planning A 22 (9) 1259–1266.

2. Bleys, B., 2008. Proposed changes to the Index of Sustainable Economic Welfare: An application to Belgium. Ecological Economics 64 (4) 741-751.

3. Brennan, A. J., 2008. Theoretical foundations of sustainable economic welfare indicators — ISEW and political economy of the disembedded system. Ecological Economics 67 (1) 1-19.

4. Brown, L. R. 1982. A Sustainable Scociety. Environment 24 (1) 2-3.

5. Bührs, T., 2004. Sharing environmental space: the role of law, economics and politics. Journal of Environmental Planning & Management 47 (3) 428-447.

6. Bührs, T., 2009. Environmental Space as a Basis for Legitimating Global Governance of Environmental Limits. Global Environmental Politics 9 (4) 111-135.

7. Collins, A., Flynn, A., 2007. Engaging with the Ecological Footprint as a Decision-Making Tool: Process and Responses. Local Environment 12 (3) 295-312.

8. Daily, G. C., 1997. Nature's Services: Societal Dependence on Natural Ecosystems. Island Press, Washington, D.C.

9. Daly, H. E., Cobb, J.B., 1994. For the common good: redirecting the economy towards community, the environment, and a sustainable future. Beacon Press, Boston MA. (orig. 1989).

10. Diamond, J. D. 2005. Collapse: how societies choose to fail or survive. Allen Lane, London & New York.

11. Distaso, A., 2007. Well-being and/or quality of life in EU countries through a multidimensional index of sustainability. Ecological Economics 64 (1) 163-180.

12. Dietz, S., Neumayer, E., 2007. Weak and strong sustainability in the SEEA: Concepts and measurement. Ecological Economics 61 (4) 617-626.

13. Ferng, J., 2001. Using composition of land multiplier to estimate ecological footprints associated with production activity. Ecological Economics 37 (2) 159-172.

14. Ferng, J., 2002. Toward a scenario analysis framework for energy footprints. Ecological Economics 40 (1) 53-69.

15. Fiala, N., 2008. Measuring sustainability: Why the ecological footprint is bad economics and bad environmental science. Ecological Economics 67 (4) 519-525.

16. Haberl, H., 2006. The global socioeconomic energetic metabolism as a sustainability problem. Energy 31 (1) 87-99.

17. Hamilton, K., 1994. Green adjustments to GDP. Resources Policy 20 155–168.

18. Hartwick, J. M., 1977. Intergenerational equity and the investing of rents of exhaustible resources. American Economic Review 67 (5) 972–974.

19. Hartwick, J. M., 1990. Natural Resources, National Accounting and Economic Depreciation. Journal of Public Economics 43 (3) 291-304.

20. Herrea-Ulloa, A. F., Charles, A. T., Lluch-Cota, S. E., Ramirez-Aguirre, H., Hernandez-Vaquez, S., Ortega-Rubio, A. 2003. A regional scale sustainable development index: the case of Baja California Sur, Mexico. International Journal of Sustainable Development & World Ecology 10 (4) 353-360.

21. Hoh, H., Schoer, K., Siebel, S. 2002. Eco-efficiency indicators in the German Environmental Economic Accounting. Statistical Journal of the United Nations Economic Commission for Europe 19 (1/2) 41-52.

22. Huppes, G., Ishikawa, M., 2009. Eco-efficiency guiding micro-level actions towards sustainability: Ten basic steps for analysis. Ecological Economics 68 (6) 1687-1700.

23. Jacobsen, T., Adams, R. M. 1958. Salt and Silt in Ancient Mesopotamian Agriculture. Science 128 (3334) 1251-1258.

24. Janssen, M. A., Scheffer, M. 2004. Overexploitation ofRenewable Resources by Ancient Societies and the Role of Sunk-cost Effects. Ecology and Society. 9 (1) 6.

25. Kabak, O., Ulengin, F., 2007. A New Perspective for Assessing the Sustainability of Countries. Journal of Transnational Management 12 (4) 3-32.

26. Kulig, A., Kolfoort, H., Hoekstra, R. 2010. The case for the hybrid capital approach for the measurement of the welfare and sustainability. Ecological Indicators 10 (2) 118-128.

27. Lawn, P. 2003. A theoretical foundation to support the Index of Sustainable Economic Welfare (ISEW), Genuine Progress Indicator (GPI), and other related indexes. Ecological Economics 44 (1) 105-118.

28. Lawn, P. 2007. A Stock-Take of Green National Accounting Initiatives. Social Indicators Research 80 (2) 427-460.

29. Lenzen, M., Murray, S. A., 2001. A modified ecological footprint method and its application to Australia. Ecological Economics 37 (2) 229-255.

30. Lozano, R., Huisingh, D., 2011. Inter-linking issues and dimensions in sustainability reporting. Journal of Cleaner Production 19 (2-3) 99-107.

31. Mickwitz, P., Melanen, M., Rosenstrom, U., Seppala, J., 2006. Regional eco-efficiency indicators – a participatory approach. Journal of Cleaner Production 14 (18) 1603-1611.

32. Mittler, D., 1999. Environmental Space and Barriers to Local Sustainability: Evidence from Edinburgh, Scotland. Local Environment 4 (3) 353-365.

33. Montgomery, R., Sanches, L., 2002. Efficiency: The sustainability criterion that provides useful guidance for statistical research. Statistical Journal of the United Nations Economic Commission for Europe 19 (1/2) 29-40.

34. Mota, R. P., Domingos, T., Martins, V., 2010. Analysis of genuine saving and potential green net national income: Portugal, 1990-2005. Ecological Economics 69 (10) 1934-1942.

35. Muller, C. C., 2008. Sustainable development: conceptualizations and measurement. Brazilian Journal of Political Economy 28 (2) 207-225.

36. Neumayer, E., 1999. The ISEW – not and indicator of sustainable economic welfare. Social Indicators Research 48 (1) 77-101.

37. Neumayer, E., 2003. Weak versus Strong Sustainability: Exploring the Limits of Two Opposing Paradigms. Edward Elgar, Northampton, MA.

38. Nordhaus & Tobin, 1972. Is Growth Obsolete? National Bureau of Economic Research.

39. Onuma, A., 1999. Sustainable consumption, sustainable development, and green net national product. Environmental Economics & Policy Studies 2 (3) 187-198.

40. Pearce, D. W., Atkinson, G. D., 1993. Capital theory and the measurement of sustainable development: an indicator of "weak" sustainability. Ecological Economics 8 (2) 103-108.

41. Peng, C., Liu, J., Dang, Q., Zhou, X., Apps, M., 2002. Developing carbon-based ecological indicators to monitor sustainability of Ontario's forests. Ecological Indicators 1 (4) 235-247.

42. Phillis, Y. A., Andriantiatsaholiniaina, L. A., 2001. Sustainability: an ill-defined concept and its assessment using fuzzy logic. Ecological Economics 37 (3) 435-456.

43. Pillarisetti, J. R., 2005. The World Bank's 'genuine savings' measure and sustainability. Ecological Economics 55 (4) 599-609.

44. Pulselli, F. M., Ciampalini, F., Tiezzi, E., Zappia, C., 2006. The index of sustainable economic welfare (ISEW) for a local authority: A case study in Italy. Ecological Economics 60 (1) 271-281.

45. Rees and Wackernagel, 1994 – Our Ecological Footprint: Reducing Human Impact on the Earth. New Society Press.

46. Riberio, R. A., 1996. Fuzzy multi attribute decision making: A review and new preference elicitation techniques. Fuzzy Sets and Systems 78 (1996) 155-181.

47. Seppala, J., Melanen, M., Maenpaa, I., Koskela, S., Tenhunen, J., Hiltunen., 2005. How can the eco-efficiency of a region be measured and monitored?. Journal of Industrial Ecology 9 (4) 117-130.

48. Shi, C., Hutchinson, S. M., Xu, S., 2004. Evaluation of coastal zone sustainability: an integrated approach applied in Shanghai Municipality and Chong Ming Island. Journal of Environmental Mangement 71 (4) 335-344.

49. Shmelev, S. E., Rodriguez-Labajos, B., 2009. Dynamic multidimensional assessment of sustainability at the macro level: The case of Austria. Ecological Economics 68 (10) 2560-2573.

50. Siche, R., Pereira, L., Agostinho, F., Ortega, E., 2010a. Convergence of ecological footprint and emergy analysis as a sustainability indicator of countries: Peru as case study. Communications in Nonlinear Science & Numerical Simulation 15 (10) 3182-3192.

51. Siche, R., Agostinho, F., Ortega, E., 2010b. Emergy Net Primary Production (ENPP) as basis for calculation of Ecological Footprint. Ecological Indicators 10 (2) 475-483.

52. Spangenberg, J. H., 2002. Environmental space and the prism of sustainability: frameworks for indicators measuring sustainable development. Ecological Indicators 2 (3) 295-310.

53. Tasser, E., Sternbach, E., Tappiener, U., 2008. Biodiversity indicators for sustainability monitoring at a municipality level: An example of implementation in an alpine region. Ecological Indicators 8 (2008) 204-223.

54. Vitousek, P. M., Ehrlich, P. R., Ehrlich, A. H., Matson, P. A., 1986. Human

Appropriation of the Products of Photosynthesis. Bioscience 36 (6) 368-373.

55. Van den Bergh, J. C. J. M., Verbruggen, H., 1999. Spatial sustainability, trade and indicators: an evaluation of the 'ecological footprint'. Ecological Economics 29 (1) 61-72.

56. Victor, P. A., 1991. Indicators of sustainable development: some issues from capital theory. Ecological Economics 4 191-214.

57. Vucetich, J. A., Nelson, M. P., 2010. Sustainability: Virtuous or Vuglar?. Bioscience 60 (7) 539-544.

58. WBCSD, 1996 - Eco-efficiency and Cleaner Production - Charting the Course to Sustainability. United Nations Environmental Programme.

59. WBCSD, 2000 – Eco-efficiency – creating more value with less impact. United Nations Environmental Programme.

60. WCED, 1987. Our Common Future. Oxford University Press. Oxford.

61. Weidmann, T., Minx, J., Barrett, J., Wackernagel, M., 2006. Allocating ecological footprints to final consumption categories with input-output analysis. Ecological Economics 56 (1) 28-48.

62. Wilson, M. D., Moffatt, I., 1994. An index of sustainable economic welfare for Scotland, 1980-1991. International Journal of Sustainable Development & World Ecology 1 (4) 264.

63. Zhao, S., Li, Z., Li, W., 2005. A modified method of ecological footprint calculation and its application. Ecological Modeling 185 (1) 65-75.