A Systems Framework of Big Data for Analysis of Policy and Strategy

Peter J Coombes
Urban Water Cycle Solutions, Carrington, NSW, Australia
E-mail: thecoombes@bigpond.com

Michael Barry
BMT WBM, Spring Hill, Brisbane, Australia
E-mail: Michael.barry@bmtwbm.com.au

The Systems Framework is discussed in many publications focused on describing projects or policies. This paper provides an overview of the Systems Framework methodologies for analysis of policy, strategy and design developed over the last decade. The framework integrates water cycle, environmental and economic processes from the “bottom up” using all available data and integrating spatial and temporal scales of behavior. Advances in computing power allowed this quantum process to be underpinned by continuous simulation of local behaviours and Monte Carlo methods. This expansionist approach to analysis reveals hidden challenges and opportunities for urban areas. The Systems Framework can be reliably and robustly applied to detailed and targeted ‘what if’ analyses, including assessments of future water security and economics under a range of climatic and population growth scenarios, and future alternative strategies or policies. The spatial and temporal detail within the Systems Framework allowed understanding, reproduction and testing of the complex interactions between waterways, reservoirs, operations, water demands, water restrictions, energy demands and financial impacts. This methodology includes hind casting of the water cycle and linked economic simulations across historical periods with known financial and resources information. The authors are developing open source and web-enabled applications that will allow greater interaction with the System Framework. This initiative is supported by a number of new science publications.

1. INTRODUCTION

The Systems Framework methodology was initially developed to understand the whole of systems performance of combinations of local, precinct and centralised solutions from the perspective of environmental and anthropogenic water systems (Coombes, 2002). It was designed to answer some of the most pressing questions of Water Sensitive Urban Design (WSUD) – and specifically to provide a defensible ‘bottom up’ method of investigating the whole of society and system benefits of distributed solutions (Coombes et al., 2002; Coombes, 2005). This methodology evolved to investigations that informed public policy in many jurisdictions including Melbourne (Coombes & Bonacci Water, 2012) and Ballarat (Coombes & Barry, 2014). Ultimately, the Systems Framework incorporated the entire water cycle, environmental and economic processes to frame policies from the “bottom up” using all available data and integrating spatial and temporal scales of behaviour. Advances in computing power have permitted this quantum process to be driven by continuous simulation of local behaviors and Monte Carlo generation of multiple replicates. This move away from “top down” reductionist methods to expansionist or systems approaches of analysis has revealed challenges and opportunities for urban areas that were hitherto obscured by other more generic and average-driven analysis techniques. A key output from this analysis is a probabilistic understanding of the changes in requirements for infrastructure and for risk management in response to combined strategies.

Development of the methodology has evolved in response to a range of influences and insights. It was essential to capture the interaction of human decisions on the physical characteristics of earth systems as outlined by Lovelock (1979). This complex system includes the earth’s biosphere, atmosphere, oceans, waterways and soil profiles in a feedback or cybernetic system. Natural systems are altered by human intervention as a function of the connectedness of all things including human behavior (Carson, 1962). Our Earth system is also responds to the cumulative impacts of exponential growth of human settlements dependent on ecosystem services or natural resources (Carson, 1962; Meadows et al., 1992). Earth systems and human welfare are also impacted by climate processes (IPCC; 2014). Ecological systems contribute to human welfare, both directly and indirectly to represent
substantial economic value (Constanza et al., 1997). Systems analysis of the timelines of growth and associated behaviours of cities allows understanding of the dynamics of urban settlements and linkages or feedback with global systems (Forrester, 1969; 1971). This allows examination of the long term drivers and impacts of population growth, industrial capital, food production, resource consumption and pollution on the future state of a region (Meadows et al., 1992).

Systems analysis is a framework of approaches and habits that is based on the behaviour of component parts of a system that are revealed by the context of relationships with other components of the system (Forrester, 1969; Meadows et al., 1992). More recent influences on the systems methodology included ideas about asset replacement or renewal strategies by Clarke (1990), network linear programming for water supply headworks by Kuczera (1993) and the Urban Water Cycle by Mitchell et al. (2001). A systems analysis understands the linked, cumulative and cyclical nature of systems in response to distributed behaviors (Coombes, 2002).

Research into “Big Data” emerged from computer sciences and systems research in the 1970s (Halievi and Moed, 2012). Big data Analysis is the term for a collection of large and complex datasets that are difficult to process or understand using traditional database management tools or data processing applications (Coombes and Barry, 2014a). Systems analysis frameworks consider the linkages and interactions between elements that comprise an entire system. A society (a market) is a system that involves a large number of transactions or decisions (behaviours) that occur at multiple scales. The systems framework aims to replicate the multiple dependent transactions throughout a system (for example with respect to intersection of a water cycle with town planning processes) – and understand or value a system that changes from the smallest distributed scales (from the bottom up) in response to the choices of people or firms in the system. Recent investigations using the Systems Framework by Coombes & Barry (2012) established that the conventional use of global averages in simulation of regional water systems is unlikely to describe the spatial and temporal contribution provided by WSUD approaches that generate distributed water resources or reductions in water demands within a metropolis - the full potential of alternative water management options including WSUD approaches may not be understood.

Aspects of this systems approach have been presented in a large number of publications dedicated to describing a project or policy. In contrast, this paper is focused on presenting an overview of the Systems Framework methodology itself, which has been developed over 20 years of applied research and practical application, and has provided insight into policies and strategies for water cycle management, urban planning, environmental protection, energy uses and economics both domestically and overseas. The methodology has been employed to provide water cycle systems performance, strategic and policy advice for national, state and local governments. The approach has also been used to design over 120 sustainable projects including city scale developments. Extension of the system principles and techniques to energy networks, open source capability and GIS environments is a topic of current research and development.

2. OVERVIEW

The Systems Framework incorporates the local scale (the people) inputs (a “bottom up” process) as a fundamental element of the method as a contrast to contemporary “top down” assumptions. Analysis is constructed from the basic elements (local land uses) that drive system behaviours and account for distributed first principles transactions to allow simulation of spatial and temporal performance of the system. Biophysical systems for a region are constructed using four basic components:

- Demands – local requirement for services and amenity
- Sources - Regional and local water sources, catchments and waterways
- Flux – transport and treatment of water, sewage and stormwater throughout the region
- Sinks – Stormwater runoff and wastewater disposal to waterways

This structure is anchored on detailed “big data” inputs, such as demographic profiles, topography and climate, and linked systems that account for water demands, water supply, sewerage flows, stormwater runoff, water quality, human health, and environmental considerations. The Framework is a series of applications for continuous simulation of water balances that interact to span all relevant spatial and temporal scales including household or land use to city to national and global scales at
timelines of one second to 100 years. The process includes multiple replicates of climate sequences and linked responses that yield probabilistic understanding of behavioral and risks. This includes water use and linked generation of wastewater, and stormwater runoff at the local scale, distribution infrastructure and information at the sub-regional or precinct scale, and regional behaviors and infrastructure such as water extractions from dams and discharges of sewage to wastewater treatment plants and ultimately to environmental receiving waters. An overview of the linked scales utilized in the Systems Framework is presented in Figure 1. The Framework is currently being redeveloped to operate in an open source GIS environment. A general overview of the hierarchy that corresponds to the conceptual description of the Systems Framework is presented in Figure 2.
Figure 1 demonstrates the linked scales that are underpinned by Big Data that are utilized in the Systems Framework. This process allows simulation of linked flows of water, nutrients, finances, sediments and energy throughout a city or a region. These processes range from the details of household behavior and associated water balances (at time resolutions of seconds) to long term forecasting of bulk infrastructure requirements or flood risks or government policies. Figure 2 reveals that the scales of analysis are linked by a hierarchy of processes that are modified by feedback loops. For example, the behavioral water demands at the local scale are impacted by water restrictions applied at the catchment scale, and climate and economic processes from the regional scale.

3. BIG DATA INPUTS

The Systems Framework includes and links multiple layers of temporal and spatial data or information from many different disciplines and from multiple perspectives. These Big Data Analysis processes are used to link and process information across space and time within a geographical information system (GIS). Big Data Analysis is the term for a collection of large and complex datasets that are difficult to process or understand using traditional database management tools or data processing applications. Most parameters that describe the characteristics and behaviour of cities are subject to significant spatial and temporal variation that is not often considered in the development of water policies (Coombes & Barry, 2012; 2014). For example, water demand is dependent on demographic, climate and socio-economic parameters that vary across a city. Considerable spatial and temporal variation in climate, stormwater runoff and water use behaviours are also observed throughout urban regions. The Big Data layer includes a unique range of spatial and temporal data processing applications that facilitate continuous simulation of system behaviors and detailed analysis.

One of these components is the development of robust and continuous meteorological data as key inputs to the analysis. This process utilizes the non-parametric nearest neighborhood schemes outlined by Coombes (2004) to transform typically inconsistent and fragmented Bureau of Meteorology data into temporally and spatially continuous surfaces of climate data. An example of a continuous rainfall surface is presented for South East Queensland in Figure 4 with associated topography in Figure 3.

Information about major water distribution, stormwater and sewage infrastructure is combined with demographic, climate and topographic data to define the zones used to compile inputs to the systems analysis. This processes also incorporate technical reports and planning documents such as (but not limited to) ABS population dynamics, water supply infrastructure, water billing records, land use, hydrology and hydrogeological records, rainfall and climate.
Land use statistics for each chosen zone are sourced from the census of population and housing. The ten categories of land uses employed as inputs to the Systems Framework are Agricultural, Commercial, Education, Medical, Industrial, Irrigated Parkland, Non-Irrigated Parkland, Residential, Transport and Water. Each of these categories can be derived using digital boundaries from the ABS, cadastral boundaries and town planning schemes. Aerial imagery is also utilised to rationalise the determination of land uses. Examples of the determination of land uses for selected state suburbs in Victoria are provided in Figures 5 and 6.

4. LOCAL SCALE

The Local Scale processes in the Systems Framework are utilised to capture the distributed local behaviours of people, buildings and land uses that drive the performance of entire system in response to the layers of Big Data inputs. A schematic of selected processes from the PURRS (Coombes, 2006) engine that continuously simulates local water balances is provided in Figure 7.

Figure 7 highlights that the local scale process utilises climate inputs (rainfall, temperature and potential evaporation) to simulate demands of mains water, recycled water and electricity, soil moisture processes, financial processes, health risks, stormwater runoff and wastewater discharges.
This process can include trees and vegetation, rainwater harvesting, onsite wastewater treatment and reuse, water efficient appliances, infiltration trenches, rain gardens and onsite detention. The local scale includes three dwelling types (detached, semi-detached and units) and five household sizes (1 person to 5+ people). These residential land uses are combined with non-residential land uses including agricultural, commerce, industry, education, medical, forests, irrigated parks and transport. Inputs to local scale simulations from the Big Data layers include demographic data (such as profiles of dwelling types and household sizes), socio-economic information (including household income, water and energy use, and business statistics), geological data (such as soil types and profiles) and climate data (rainfall, temperature and evaporation).

Local scale continuous simulations are completed for each dwelling type and land use throughout the entire region at time steps ranging from one second to six minutes using the longest local sequences of rainfall data. The local scale simulations include development of behavioural indoor and outdoor water demands (Coombes et al., 2000; Cui et al., 2008) and the distribution of indoor end uses that are calibrated to local observed data. For example, the distribution of indoor end uses for a suburb in Sydney and calibration to residential billing records is presented in Figures 8 and 9, respectively.

Whenever possible, the calibration is conducted in a “base year(s)” period that is free of water restrictions and other unusual events (such as floods or major economic processes) that would affect “base” water use behaviours. The external impacts of water restrictions and other unusual events on normal operation of local water balances are applied from the transition, catchment and regional scales in the Systems Framework. Generated sequences of stormwater runoff and energy demands are also calibrated to any available local data. Outputs from the Local Scale analysis include long sequences of water demands, wastewater discharges, stormwater runoff, energy demand, water quality, soil moisture and finances. These results are combined with climate data and passed to the Transition Framework as reference files.

5. TRANSITION SCALE PROCESSES

The sequences of water use, wastewater flows, stormwater runoff, financial transactions and energy use from the local scale analysis are combined for each zone or sub-catchment using town planning projections and replicates of spatial climate sequences (Figure 10). Zones or sub-catchments are chosen as the intersection of unique zones for town planning, demographics, topography, climate, inflows to waterways and infrastructure, and water supply. This process combines local scale Options to produce future projections of water demands, wastewater flows, stormwater runoff, energy use and financial processes for use in the Systems framework (Figure 11). These sequences are then verified and calibrated using hind casting of historical inputs for a period of observed behaviour. This process utilises behavioural calibration parameters to capture the unique spatial and temporal transition of local scale results to the chosen zone scale.

A transition framework is used to generate daily water cycle responses for each zone. Sequences of daily water and financial balance results from local scale are linked using seasonal information and
historical climate data (including daily rain depths, cumulative days without rainfall and average daily maximum ambient air temperature) to create resource files of water demand, wastewater generation, stormwater runoff, energy use and economic transactions. This method of non-parametric aggregation (Coombes et al., 2002) generates daily outputs from each zone using the historical resource files and climate replicates generated for the simulation of the regional system. Climate replicates are multiple sequences of equally likely future climate drivers (such as rainfall, temperature) that are generated using Monte Carlo processes.

Figure 10 demonstrates that at each time step climate variables from the regional model are used to find matching climate variables and coincident daily water use, wastewater generation, stormwater runoff, and financial results for each dwelling and all land uses from the reference files. These results are combined with population, town planning and demographic data at each time step to estimate total indoor and outdoor use, sewage flows and stormwater runoff for each zone. The sequences of data from the local simulations were combined in the transition framework using the process presented in Figure 11. Daily sequences of water cycle information; such as water demands, wastewater discharges and stormwater runoff; are combined for different household sizes and land uses, different dwelling types and a combination of different water cycle management Options for each strategy in the Transition Framework. The climate variables from the Big Data layer are also used in the regional systems model and are derived using the synthetic climate series generated using historical climate sequences. Importantly the climate replicates are temporally and spatially consistent with the rainfall and stream flows in the water supply catchments to maintain direct behavioural links between local water use and regional catchment behaviours.

6. CATCHMENT AND REGIONAL SCALE PROCESSES

The Systems Framework combines water, wastewater and stormwater infrastructure networks with waterways and catchments in an integrated network. Spatial and temporal information generated by the lot scale simulations are combined by the zone scale transition as inputs to the network analysis. The movement of water, wastewater, recycled water, energy, finances and stormwater throughout a region is simulated over a chosen planning horizon using multiple replicates of climate sequences generated using Monte Carlo processes. This simulation of integrated networks and catchments allows analysis of the probability of peak flows in trunk infrastructure, assessment for regional wastewater discharges, stormwater runoff, water demands, stream flows, energy use, economics, flooding and pollutant loads.

The regional scale of the Framework includes water sources from ground water, surface water sourced from regional river basins, shared surface water with other river basins, wastewater reuse and
stormwater harvesting. The linked analysis utilizes stream flows, reservoir storage volumes, wastewater discharges, information about the operation of water systems and data from global climate model as inputs. An example of the water resources network for Greater Sydney is presented in Figure 12.

Figure 12: Water resources layer (rivers, water supply, distribution and demand) for Greater Sydney

Figure 12 demonstrates that regional scale links regional water resources layers to the zones that represent the combined behaviour of each local government area. Streamflow in water supply catchments is generated using the SIMHYD algorithms described by Chiew et al., (2002) with inputs from the Big Data surfaces of topography and climate. The Systems Framework optimises the water balances at each node using network linear programming, user defined system objectives and boundary conditions in accordance with the principles outlined by Kuczera, (1993). For example, the framework can be required to meet of global objective of ensuring all water demands are supplied and minimum flows in waterways are maintained within the capacity constraints of water infrastructure. In addition, use of game theory concepts from Nash, (1950) relating to competitive games - best response where each player in a game selects the best response to the other players’ known best strategies - allows optimization of multiple transactions or behaviors at the scales feeding the regional analysis. The simulations also include operating rules and regional policies such as water restriction triggers. The behaviour of the System Framework is verified at the regional scale using a hindcasting process that compares predicted and observed behaviors for key processes within historical time periods. For example, the hindcasting results for water demands for Greater Sydney and streamflows in the Shoalhaven River are provided in Figures 13 and 14, respectively.

Figure 13: Hindcasting of water demands for Greater Sydney

Figure 14: Hindcasting of streamflow for the Shoalhaven River
Note that verification of water demands also confirms the accuracy of the water restriction rules and local water use behaviors. The linked multiple scale processes of water demands, stormwater runoff and wastewater discharges can also be verified at the regional scale by the hindcasting to observed wastewater inflows to wastewater treatment plants (Figure 15 demonstrates comparison at the Western Treatment Plant in Melbourne) and to observed flows at stream gauging stations. Similarly, the distributed and cumulative costs of providing, replacing and operating infrastructure are compiled at the regional allowing unique hindcasting verification of finances as shown for Ballarat in Figure 16.

The use of multiple replicates of climate allows understanding of the frequency, percentiles and probability of the behaviours of different options across a region. For example, water demands and wastewater discharges for different Options (Conventional, PER: stormwater and rainwater harvesting, Flow Systems: wastewater reuse) in the Caloundra region is presented in Figure 17.
Figure 17 reveals that PER Option with water efficient appliances, rainwater and stormwater harvesting, and the Flow Systems Option with wastewater reuse creates substantial changes in the distributions of daily water demands and wastewater flows. This type of probabilistic analysis is a primary benefit of the Systems Framework approach which provides understanding of the changes in requirements for infrastructure and for risk management in response to combined strategies.

The Systems Framework also includes financial transactions at all scales to account for the costs and greenhouse gas emissions of providing services, and the revenue generated by those services. For example, the distributed net present values and greenhouse gas emissions are presented in Figures 18 and 19, respectively.
Figures 18 and 19 demonstrate that the structure of the Systems Framework allows analysis of the spatial and cumulative nature of water cycle systems. Analysis of linked scales and water cycle allows understanding of the distribution of costs across different jurisdictions (such as local government, water authorities, the environment and building owners) and the trade-offs between elements of the water cycle (including waterways, water supply, soil profiles and stormwater runoff) created by alternative Options. The Systems Framework facilitates analysis of the linked economic responses or the value chain of households, state and local governments, bulk and retail water authorities, and the private sector.

7. INSIGHTS

This paper provided an overview of the Systems Framework methodologies and capability for analysis of policy, strategy and design that were developed over the last decade of investigations underpinning development of public policy. A Systems Framework for Big Data analysis was built on pioneering systems investigations of policies for economic development that commenced with Urban Dynamics (Forester, 1969), Earth Dynamics (Forester, 1971), Limits to Growth (Meadows et al., 1992) and, more recently, Asset Replacement (Clarke, 1990), Network Linear Programming for Water Supply Headworks (Kuczera, 1993) and the Urban Water Cycle (Mitchell et al., 2001). The systems philosophy was expanded to further incorporate the resilience and connectedness of natural systems, human behavior, infrastructure networks and climate processes. Ultimately, the Systems Framework incorporated the entire water cycle, environmental and economic processes to frame evidence based policy from the “bottom up” using all available data and integrating all the spatial and temporal scales of behavior.

Advances in computing power have permitted this quantum process to be driven by continuous simulation of local behaviors and Monte Carlo generation of multiple replicates. This move away from reductionist methods to expansionist approaches of analysis has revealed a range of challenges and opportunities for urban areas that were hitherto obscured by more generalised analysis techniques. This methodology includes the full cumulative costs (and benefits) of projects for water cycle management across an entire system and allows greater understanding of the trade-offs and benefits throughout the water cycle, environment and society. This methodology captures the dynamics of urban systems that operate at multiple spatial, temporal and dimensional scales to generate water demands, wastewater flows and stormwater runoff, costs of transferring and providing services. It is a key insight that centralised system behaviours are highly sensitive to variations in local policies, strategies, solutions and behaviors, and respond in a cumulative (not linear or static) manner. The value of distributed policies and actions (such as WSUD) can only fully realised from a systems perspective.

The Systems Framework can be reliably and robustly applied to detailed and targeted “what if” analyses, including assessments of future water security and economics under a range of climatic and population growth scenarios, and future alternative strategies or policies. The spatial and temporal detail within the Systems Framework allowed understanding, reproduction and testing of the complex interactions between waterways, reservoirs, operations, water demands, water restrictions, energy demands and financial impacts. This methodology includes hind casting of the water cycle and linked economic simulations across historical periods with known financial information about the costs of the providing services in the region. This process can be used to verify and calibrate the inputs to and processes within the water resources and economic framework. In general, the hind casting process was able to verify the efficacy of the system behaviours and economics. Finally, the authors are developing open source and web-enabled applications that will allow greater interaction with the System Framework. This initiative is supported by a number of new science publications.

8. REFERENCES


IPCC., (2014). Climate change 2013: the physical science basis. Fifth assessment report of the Intergovernmental Panel on Climate Change.


