6

MULTIPLE ACTORS AND CONFOUNDING FACTORS

Evaluating impact in complex social-ecological systems

Aaron E. Zazueta and Jeneen R. Garcia

The GEF and its support in the South China Sea¹

The GEF is a financial mechanism that provides grant and concessional funds to 183 participant countries for projects and activities that help the countries fulfil their commitments to global environmental conventions.² From its inception in 1992 to the time the evaluation began in 2010, GEF support for coastal and marine-related concerns in the South China Sea was delivered through 34 GEF projects (\$112 million) and 150 small grants (\$3 million). This support was provided to seven littoral countries that were eligible for GEF funding – Cambodia, China, Indonesia, Malaysia, the Philippines, Thailand and Viet Nam. The projects were implemented at different times by various GEF partner agencies, which include the Asian Development Bank, United Nations Development Programme, UN Environment (formerly UN Environmental Programme) and the World Bank.

The main objective of the impact evaluation was to assess the extent to which GEF support contributed to, or is likely to lead to, changes in policies, technology, management practices or other behaviours that will address the priority trans-boundary environmental concerns.³

The geographical focus of the evaluation included the Gulf of Tonkin and Gulf of Thailand as part of the South China Sea. The marine area covered by the evaluation stretches in a south-west to north-east direction with a southern border between South Sumatra and Kalimantan (Karimata Strait), and a northern border at the Strait of Taiwan between the northern tip of Taiwan and the Fujian coast of mainland China.

The South China Sea – as is the case with other, large marine ecosystems – is characterized as a complex social-ecological system due to its inextricably linked and constantly interacting natural and human systems (Anderies 2004; Liu et al. 2007; Ostrom 2009). As a complex system, any interventions in the South China Sea

are subject to non-linear processes that often make prediction of outcomes difficult (Holling 2001; Taleb 2007).

First, the time scale at which interventions take place (typically five to seven years) is often disproportionate to the time scale at which the ecosystem responds. Ecosystem responses may be immediate or may span one to several decades. Second, the geographical scale at which interventions take place (typically a local site, municipality or country) is likewise often disproportionate to the geographical scale of the ecosystem boundaries, or the boundaries within which the environmental concern being addressed originates and extends its impact. Ecosystem boundaries often transcend municipal or country boundaries, while environmental concerns may cut across economic and social system boundaries. Interventions implemented at a local scale may be affected by events and other factors interacting at the national, regional or global scales, such as market forces.⁴ Third, the factors and actors at play influence each other and therefore make it difficult to trace the chains of causality from the observed change back to the intervention being assessed. Even if changes can be ascertained, and links between any given intervention and an observed change can be made, multiple actors and other confounding factors make it difficult to attribute causality to any specific intervention (Sanderson 2000; Davies 2004; Mee, Dublin and Eberhard 2006).

Thus, impact evaluation in complex systems requires close attention to the ways in which an intervention interacts with other elements of the system to influence the observed changes. To understand the impact of GEF support in the South China Sea, the evaluation drew on scientific literature on complex systems to guide the lines of inquiry that would (1) help assess the extent of change in environmental conditions, (2) establish causal links between interventions and the observed changes and (3) identify the specific contributions of GEF support to these observed changes within the context of other actors' contributions, and the influences of confounding factors.

Assessing direct impact

Determining the extent of change

Determining the extent of change that took place since GEF support began was in some instances straightforward. In the case of the Livestock Waste Management Project in China, Thailand and Viet Nam (GEF ID 2138) that demonstrated new pollution control technologies, end-of-pipeline measurements before and after the introduction of the technology provided reliable measurements of nutrient loads prevented by the intervention. Other indicators of change were improvements in the smell in pig farms and reductions in incidents of red-eye disease among farmers in all three visited demonstration sites. In Ha Tay, Viet Nam, evidence of pollution reduction was that the wastewater could be used for agriculture – wastewater released into rice paddies no longer killed the plants.

In the case of mangrove ecosystem management interventions, one important indicator was change in mangrove cover. Changes in mangrove cover were

determined from before-and-after data presented in terminal evaluations and annual project reports. Data were verified through field visits that included visual assessments of mangrove forest conditions and interviews with stakeholders on the extent to which changes had taken place. To fill in information gaps and triangulate evidence from these information sources, the evaluation team carried out remote sensing analysis to compare mangrove cover at the sites across different years. Evidence confirmed that in most sites where the GEF had provided support, mangrove cover had increased over a period of about five years.

Given the nature of the concerns that GEF-supported interventions address, time lags of such magnitude (i.e. five years) or greater are necessary before changes in environmental status become evident. The higher the scale⁵ of the system being addressed, the more slowly it tends to respond and, consequently, the longer it takes to begin shifting to an improved social and environmental state (Gunderson and Holling 2002). Further, several interventions may need to take place before a tipping point is reached, triggering a cascade of effects in the social-ecological system (Gladwell 2000; Kinzig et al. 2006). These unpredictable time lags between interventions and responses in the social-ecological system require that impact evaluation not only focus on changes that are directly observable in the social-ecological system (i.e. income of the local population or pollution levels) but also on indicators or markers of future change that allow an assessment of the extent to which interventions are likely to lead to the desired impact.

For example, numerous studies have shown that coral reef ecosystems may take decades to centuries before the effects of overfishing are observed (Jackson et al. 2001; Folke et al. 2004; Dulvy, Freckleton and Polunin 2004; Babcock et al. 2010). The evaluation found that while coral cover had been maintained or was slightly increasing in GEF-supported marine protected areas (often referred to as MPAs) over the course of at least five years, larger fish species and giant clam abundance had decreased in the Con Dao and Phu Quoc marine protected areas in Viet Nam. Although evidence indicates that marine protected area initiatives have positive results for hard corals, scientific research shows that declining fish populations may soon be followed by coral reef decline at the ecosystem level (Pandolfi et al. 2005). Thus, to better assess the likelihood of an intervention's impact on coral reef ecosystems, indicators tracking fish and benthic organisms (bottom-dwellers) must be measured. On this basis, the evaluation concluded that though GEF projects had often been able to protect coral cover from physical destruction, impact has been limited and insufficient to prevent the degradation of coral reefs targeted by these projects.

Establishing a chain of causality

In cases in which interventions were highly targeted and change could be measured immediately and in a straightforward manner, such as in the reduction of nutrient pollution, causality is easy to establish. However, interventions more commonly do not have a one-to-one or linear relationship with a desired effect (Stern et al.

2012). In cases where interventions were implemented at higher scales (e.g. national and regional) and sought to address non-linear processes, causal links between interventions and observed changes were often more difficult to establish.

A critical starting point for impact evaluation in such cases is to develop a sound understanding of the human and natural dynamics inherent to each of these interacting systems. This understanding enables the evaluation to identify the geographical boundaries of the specific system in which the intervention needs to create an impact; the system's inherent characteristics defining the extent of change that can take place, and the types of interventions that are most likely to result in change, given the patterns of causality in the specific system.

Mismatches between the scale at which the intervention is implemented and the scale at which the environmental concern manifests are common when the system's geographical boundaries are not well understood. This has often resulted in administrative or management boundaries not being aligned with ecosystem boundaries (Folke et al. 1997, 2002; Cash et al. 2006). For example, migratory fish stocks typically transcend national boundaries, requiring the countries concerned to agree on and create management boundaries that better approximate that of the fish stocks.

In the South China Sea, the evaluation found that while countries were aware of this, they historically have tended to not enter into binding multilateral agreements, especially when it came to environmental concerns. Of the 27 multilateral and bilateral regional instruments pertaining to coastal management in force in the South China Sea, only four bilateral agreements were legally binding. A comparative assessment of regional mechanisms in similar large marine ecosystems around the world also showed that even after more than 20 years, most regional mechanisms still relied heavily on external donors when low- to middle-income countries comprised the majority of the membership.⁶

Considering the countries' unwillingness to enter into binding regional environmental agreements, the relatively short time frame and the wide differences in country incomes, GEF support in the South China Sea region could not be expected to facilitate the ratification of a regional convention such as those present in other regional seas.

Evidence from scholarly research, however, indicates that in systems in which no single central actor disproportionately influences the behaviour of others (such as the South China Sea region), collaborative adaptive behaviour among independent actors tends to be more prevalent where good communication exists, especially when collaborative behaviour is expected across multiple administrative or ecological scales (Brondizio, Ostrom and Young 2009; Ostrom 2010). Therefore, one way to measure GEF's contribution towards regional environmental governance was to assess the extent to which GEF support fostered better communication, and the extent to which indicators of cooperative behaviour among countries could be linked to GEF support.

GEF-supported regional interventions included multidisciplinary workshops, steering committee meetings, task forces and expert networks that provided

many new opportunities for interaction across the region among countries, local governments, regional organizations, civil society (including academic and research institutions) and the private sector. These activities allowed information exchange, joint fact-finding and regional priority setting and planning. Social network analyses of regional environmental actors carried out by the evaluation team based on available reports, online surveys and key informant interviews revealed that GEF support helped increase the connectivity between national and regional actors.

Furthermore, the evaluation team used interviews and document analyses to establish that new collaborative agreements were adopted as a direct result of GEF projects. This finding further strengthened evidence of causal links between GEF-supported communication-related activities and collaborative adaptive behaviour among independent actors, which in the case of the South China Sea is the most critical outcome given the region's inherent characteristics. Among these collaborative agreements are the memorandum of agreement between the provinces of Kien Giang (Viet Nam) and Kampot (Cambodia) signed in 2008 to jointly manage trans-boundary seagrass beds; the engagement of all littoral countries, including China, in the elaboration and approval of a Strategic Action Programme for the South China Sea; the Gulf of Thailand Joint Statement and Framework Programme for Joint Oil Spill Preparedness and Response, signed by Cambodia, Thailand and Viet Nam and supported by the private sector and international organizations; and the adoption of the Sustainable Development Strategy for the Seas of East Asia (SDS-SEA) by 11 East Asian countries and 19 non-country partners.

Although all these actions were taken by governments and other independent actors, stakeholder interviews and document analyses clearly indicated that these collaborative agreements would have been very unlikely at the time they took place had GEF support not been present. These accomplishments are particularly significant considering the tensions that exist among the littoral countries and the slow progress that had been made in multilateral processes related to the environmental management of the South China Sea. Based on these pieces of evidence analysed within the region's contextual conditions, the evaluation concluded that the GEF had provided the appropriate type of support at the regional scale, leading to outcomes that were critical to achieving long-term impacts.

As the analysis progressed, it became increasingly clear that the direct results of GEF projects tended to be strikingly similar within and across the different scales. Thus, GEF support was found to have a somewhat recursive nature across scales in as far as projects supported three broad types of interventions, including: generating and expanding information and knowledge (including raising awareness and developing mechanisms for information sharing and access); developing institutional capacities (including policy and regulatory frameworks, administrative structures and processes for trust building); and testing implementation strategies (including introducing technologies and approaches, implementation mechanisms and bodies and financial instruments).

The evaluation also found that the GEF has provided considerable support to initiating processes that facilitate the broader adoption of implementation strategies

introduced by project demonstrations. Based on these findings, the evaluation developed a framework that was used to identify and classify project results and to establish chains of causality between GEF support and the observed changes. This framework proved to be very useful during the evaluation as a tool to compare and establish links of GEF engagement at different scales, regardless of the project type country or implementing agency.

Discounting rival hypotheses

Discounting rival hypotheses entails assessing which feasible causes or alternative explanations account for the change observed, and which do not. Most important, it establishes that the intervention being evaluated is one of the feasible causes, and that the observed change can be attributed to it. The evaluation used several approaches to discount rival hypotheses, depending largely on the extent to which there were confounding factors at play and the extent of information available. The first task was to assess if there were rival hypotheses with merit. For example, attributing nutrient pollution reduction to the new technology introduced in livestock farm demonstrations was straightforward, as the causal link was clear and field visits and interviews confirmed that no other plausible explanations could account for such a sharp decrease in nutrient loads in wastewater.

In the case of activities that supported mangrove management, a rival hypothesis with merit was that widespread policy changes could account for, or at least significantly contribute, to the change in the mangrove cover trends in the demonstration sites. Counterfactual analysis (i.e. comparison of results in similar sites where mangroves had not benefited from an intervention such as that provided by GEF projects) would have been desirable to assess the extent to which broader factors affected local trends in demonstration sites. But given the various socio-economic and environmental factors affecting mangrove growth, compounded by the lack of environmental monitoring data in both GEF and non-GEF sites, it was difficult to determine the counterfactual.

In the absence of information on counterfactuals, the evaluation compared trends in mangrove cover data at the demonstration site to more readily available data on national- and regional-scale trends. In most cases, it was found that GEF-supported sites showed some recovery, whereas national and regional trends continued to decline, albeit at a slower rate. This indicated that interventions at GEF-supported sites were indeed resulting in positive change compared to the rest of the region and countries that had not received these types of interventions. However, the trends also showed that widespread policy changes were contributing to this decline in mangrove cover slowing down across the region.

The evaluation also considered ongoing processes and other contextual factors that could have contributed to the observed changes to better assess the extent of GEF contributions. From remote-sensing analysis and interviews, the evaluation found that in most cases, mangrove cover was already increasing before GEF support began. In Trat, Thailand, for example, the process was driven by local communities

as well as national economic conditions as early as 10 years before any GEF projects took place. GEF support, however, allowed communities in conflict to develop rules for mangrove use that then encouraged cooperative protection of a greater area of the mangrove forest. The impact of GEF support then was not increasing mangrove cover per se, but sustaining and increasing the momentum of community and government support for mangrove protection that otherwise might have fizzled out with no follow-up funding during that critical period.

At the regional scale, where cases were more complex, and plausible rival hypotheses grew in number or were more entangled, the analysis of context was even more important. The evaluation team made historical analyses of actions, programmes or events that took place before, during and after GEF support. For example, when assessing the impact of GEF support on the regional environmental architecture related to marine and coastal matters, the evaluation started by determining the results of efforts to establish regional mechanisms that predated GEF support. The evaluation also assessed the ways in which previous attempts to establish regional mechanisms affected GEF-supported initiatives.

Among the contextual factors, the Rio processes related to the summits on sustainable development that have taken place in the last three decades since 1992⁷ were found to be key to raising awareness on environmental issues among policy makers, as these coincided with a cluster of regional and national environmental initiatives being implemented at specific periods. The Rio processes also implied a set of commitments to which countries signatory to the global conventions agreed to meet, often using their own resources but also drawing on resources provided by donors and multilateral financial institutions.

In the context of this multiplicity of actors and conditions affecting country decisions, it became a challenge to assess what had been the GEF's contributions to achieving desired impacts that often had not yet manifested. Thus, at the regional scale, the evaluation's approach was to determine the extent to which the GEF helped create the conditions and contributed to regional processes that in the causal chain are expected to lead to environmental stress reduction and improved environmental status.

To determine if the GEF had a unique contribution or niche in the mix of all these actors, the evaluation carried out a comparative assessment of the support provided by the GEF and other donors to marine and coastal initiatives in the region using information from a global donor database, historical records, project documents and donor reports. The idea was to place GEF support in the context of the financial support landscape in the East Asian seas over the last 25 years. The analysis found that the GEF was by far the biggest donor when it came to marine and coastal initiatives at the regional scale. Interviews and desk reviews of other regional organizations' reports also showed that the GEF was the only actor that linked its initiatives at local, national and regional scales. An important GEF contribution, this enabled actors to interact more effectively and allowed the processes necessary for environmental stress reduction in the South China Sea to progress in parallel at different scales. Furthermore, the regional mechanisms that

the GEF supported had provided a framework that linked the initiatives of other donors and actors mainly working at a national scale to existing regional efforts.

Social network analysis was used to compare the current interactions among regional actors with GEF support present to a scenario without GEF support. The evaluation found that several of the organizations that GEF supported had over the years become more central to the network of regional environmental actors. With GEF support, some of the long-active organizations were able to expand their activities to include regional-scale coastal and marine initiatives. In general, more links among organizations existed in the presence of GEF support compared to the hypothetical scenario without GEF support, where some actors were completely cut off from the regional network. This affirmed the finding that one of the GEF's important contributions to the regional environmental architecture was the linking of initiatives and actors across local, national and regional scales to address priority regional transboundary concerns.

In both of the examples on the impacts of GEF support on mangrove cover at the local scale, and the environmental governance architecture at the regional scale, the evaluation team triangulated different information sources and analytical methods in order to augment the lack of available information, and to account for confounding factors or limitations imposed by the social-ecological system's inherent characteristics. In the absence of counterfactuals, the task of impact evaluation is to develop approaches to discount rival hypotheses that, although not able to attribute changes solely to the intervention being assessed, are nevertheless able to make a clear link between the intervention's contributions and the observed outcomes (Mayne 2011).

Assessing progress towards impact

The undertaking of the three tasks of impact evaluation showed that the GEF had helped countries improve capacities and identify approaches that resulted in positive environmental changes. However, the evaluation also determined that much of the accomplishments were found to have taken place at a relatively small scale or in specific conditions in which the GEF-supported demonstrations had been implemented. Despite this progress, the gains were insufficient to address the full range of challenges in the sustainable management of the South China Sea, GEF's goal in the region.

Fully achieving impacts requires that the boundaries of the intervention match the boundaries of the system or the environmental concern being addressed. For example, the evaluation found that GEF-supported demonstrations protecting seagrass and coral reefs were often effective in addressing local-scale stressors and helped slow down rates of environmental degradation in the short term. However, the evaluation also found that GEF support during that period was less successful in addressing stressors that originated beyond these localities, such as siltation from the watersheds in which the protected areas were situated or encroachment by commercial fishers from neighbouring countries with lower environmental awareness and a lower interest in protecting the local resources.

These larger-scale drivers have inevitably resulted in continuing trends of degradation. It is not realistic to expect that the type of support provided by the GEF or other agencies can by itself counteract the numerous factors driving environmental degradation in the South China Sea, such as global demand for natural resources, population growth and increasing standards of living in the region. Larger-scale changes will have to be initiated at the scale of the countries and the broader consumer markets if positive environmental impact is to be sustained over the long term.

Thus, GEF support is intended to catalyse processes leading to these large-scale changes, with the engagement of country governments, private sector, civil society organizations, local communities and other stakeholders that, with time, will result in an improved environment (van den Berg and Todd 2011). Reaching this larger goal will, among other things, require a much broader adoption of the approaches demonstrated with GEF support. Though these processes may be initiated during a project's lifetime, such broader adoption is not expected to take place until perhaps decades after GEF support has ended.

In this context, a central concern of impact evaluation is to assess not only direct, immediate impacts, but also – and more important – the progress made towards achieving impact at the scale of the targeted system by assessing the extent to which broader adoption has begun, will continue or is likely to take place beyond project completion. Particularly important is assessing which approaches are truly effective, how broader adoption occurs, under what circumstances and why (Pawson and Tilley 1997; Pawson, Greenhalgh and Walshe 2005).

The evaluation found that broader adoption took place through three main processes. One was through mainstreaming, by which information, lessons or specific results of GEF projects were incorporated into broader stakeholder mandates and initiatives such as laws, policies, regulations and programmes. This might occur not only through governments but also within communities, other development agents and the private sector. Another process was through replication, by which GEF-supported initiatives were reproduced or adopted at comparable administrative or ecological scales, often in another geographical area or region. The third process was scaling-up, by which GEF-supported initiatives were expanded to cover larger geographical scales, often including new aspects or concerns that might be political, administrative or ecological in nature. Scaling-up allows the integration of concerns that cannot be resolved at lower scales and allows the spread of GEF contributions to areas contiguous to the original project sites.

These three processes of broader adoption may be at work in any combination at the same time for a given project and may take place simultaneously at different spatial and temporal scales, in different degrees of intensity and importance. Often, one process may have to occur for another process to take place. Assessing the effects of GEF support at a broader scale entails assessing evidence of these processes occurring as they relate to the broader adoption of technologies and approaches that have been shown effective in improving management of marine and coastal resources.

Contextual factors directly and primarily affected the differences in the extent of broader adoption and the extent or likelihood of long-term impact of GEF support. For example, although GEF support to the integrated coastal management (often referred to as ICM) approach was fairly consistent across the East Asia region and mostly appropriate for the demonstration sites that had been selected, scaling-up or replicating ICM beyond demonstration sites has greatly advanced mainly in countries such as China and the Philippines. In these countries, national policies of decentralization were more robust and local economic conditions and institutional absorptive capacities in the expansion areas were more developed. This is in contrast to processes taking place in countries such as Viet Nam or Thailand, where national decentralization policies were not as robust and where local conditions in the areas were not as conducive to ICM expansion.

Another factor at play was the extent to which the specific approach introduced was an appropriate response to the environmental concerns that were relevant to the locality. In Thailand, for example, it was found that the ICM approach is well suited for addressing the planning needs of Sriracha, a port city, as well as its adjacent coastal municipalities in Chonburi province. But it is not as well suited for dealing with livestock and agricultural waste, which is a non-point source pollution originating from the province's upland areas. Thus, although scaling-up from the demonstration in Sriracha to the entire coastal area of Chonburi was found to improve wastewater management in the area, developments in the upland areas were found to take place at a much slower rate. The evaluation therefore concluded that progress towards impact was taking place in the coastal municipalities and would likely result in large-scale impact in the long term; on the other hand, GEF support would be more effective if replicated or scaled up in similar coastal areas rather than scaling up the approach to upland areas where evidence of progress towards impact was not observed.

Conclusion

The evaluation found that in most areas where GEF support has taken place, evidence of positive direct and indirect change could be observed at the local, national and regional scales. At the local scale, in cases where evidence was available, most of the technologies and approaches demonstrated through GEF support directly contributed to reducing environmental stress and, in some cases, resulted in improvements in environmental status. At the regional scale, GEF-supported initiatives that encouraged collaboration and communication and trust-building across different scales; communication and trust serve as enabling conditions to regional environmental cooperation where several sovereign countries are involved. Most important, the triangulation of evidence showed that GEF support had a unique role in the region that linked interventions at multiple scales, which has resulted in significant contributions that no other actor could have made.

However, at the time of evaluation, these positive changes were a far cry from the impact that GEF support aims to achieve in the large marine ecosystem as a

whole. Many environmental and socio-economic drivers had not been addressed due to the highly localized nature of many of the interventions. Perhaps even more important, not enough time had elapsed for more significant changes to manifest, especially at the regional scale. Under these conditions and given the larger GEF goals, the immediate or direct impacts observed merely indicated the extent to which the supported technologies and approaches were effective in addressing a specific concern in a particular context.

To evaluate the extent to which the GEF is meeting its broader goal, it is critical to assess not only the direct impacts that can be observed but also, more important, the likely intended long-term effects and indirect impacts of an intervention. To determine the catalytic or transformational contributions of any given programme or intervention, the evaluation must thus assess the extent to which there is progress in establishing the processes and conditions that will likely result in improved environmental conditions in the larger system.

This assessment is only possible if the evaluator has a sound understanding of the system being targeted by the intervention. This includes a good understanding of the system components, system boundaries and the chains of causality through which changes in the system take place. It is based on this sound understanding that the evaluator can assess if current interventions are contributing to conditions that are likely to lead to desired changes in the distant future. Understanding the system and how the system develops over time also provides the evaluator with a sense of the extent of change that is possible within a particular context, regardless of the ambitious objectives set by projects in their proposals. Similarly, through the better understanding of the chains of causality across geographical and temporal scales, the evaluator can assess which major gaps remain in the conditions necessary for change, and which critical ones need to be addressed to increase the likelihood of impact at a broader scale.

Postscript, 2020

Nearly ten years after the evaluation was completed,⁸ the authors had the opportunity to undertake a post-completion evaluation aimed at assessing the sustainability of outcomes generated by the cluster of projects linked to the establishment of the Partnerships in Environmental Management for the Seas of East Asia (PEMSEA). PEMSEA, a regional organization that itself was an outcome of the GEF-supported cluster of projects, was envisioned as the main catalyst for the broader adoption of integrated coastal management (ICM) in the region. This unusual opportunity to revisit our evaluand allows us to verify some of our assumptions about the conditions that catalyse progress towards impact. This has also allowed us to reflect on our understanding of what evaluations can and cannot assess in interventions that aim for long-term, large-scale impact in complex systems.

Below we examine three of our key conclusions and related assumptions against our updated findings from the post-completion evaluation. Afterwards, we offer an alternative approach to evaluating interventions in large complex systems.

Initiatives to increase communication among governments and other stakeholders will foster collaborations to address regional environmental challenges. The Sustainable Development Strategy for the Seas of East Asia (SDS-SEA), intended as a regional framework and platform for collaborative action among the East Asian countries and its non-country partners, was updated in 2015 with continuing GEF support. From its original membership of 12 countries in 2003, the SDS-SEA signatories now number 11 countries, with its non-country partners increasing from 19 to 21. Included among the non-country partners are NGOs, multilateral organizations and private sector enterprises. Partners continue to meet every three years at the ministerial and technical levels through the East Asian Seas Congress; a smaller Executive Committee meets twice a year to oversee implementation of the SDS-SEA. Partner countries report on their progress during the meetings, exchange learning from their experience and discuss future directions.

The PEMSEA Network of Local Governments (PNLG), a mechanism for communication, learning and exchange among local governments that was created with GEF support in 2001 has continued to grow. Since 2011, its membership has doubled to 50 local governments in 10 countries, with a self-sustaining secretariat in Xiamen, China. The Gulf of Thailand oil spill preparedness initiative, signed by Cambodia, Thailand and Viet Nam in 2006, has evolved into an organization with a secretariat based in Thailand. The three countries have agreed on the coordinated implementation of the Gulf of Thailand Strategic Action Plan (2017–2021). Based on these developments, it appears that GEF support for communication initiatives has indeed contributed to countries continuing to work together, with the progress aided by uninterrupted grants over more than 25 years.

Broader adoption and significant environmental change at a regional scale is not expected to take place until decades after GEF support has ended. The updating of the SDS-SEA in 2015 included the adoption of a target to have 25% – previously 20% – of the coastline of participating countries under ICM by 2021. The extent and thoroughness of ICM adoption differs among countries; nonetheless, all the GEF project recipient countries except one had exceeded the 25% target by 2018.⁹

Currently all participating countries have ICM sites with intersectoral committees at the provincial level, and local staff to manage the ICM process. In total, 43 provinces across eight countries have ICM learning sites. An additional 76 sites are in various stages of ICM development and for which there is little information available.

The most aggressive adoption of ICM has taken place in China (32 sites), the Philippines (32 sites), Thailand (24 sites) and Viet Nam (14 sites). These developments challenge the evaluation's conclusion that broader adoption of ICM can only take place in highly decentralized countries such as China and the Philippines, rather than in Thailand and Viet Nam. According to country representatives interviewed for the post-completion evaluation, the ICM approach has been useful for addressing multiple intersectoral issues, and not just coastal or environmental.

In this case, it appears that time was the main constraint in the speed of ICM adoption rather than the inappropriateness of the approach.

However, technical and human resources continue to be a limitation that has slowed ICM adoption. Common across the GEF-supported countries was how PEMSEA-trained government staff eventually take on more financially attractive career opportunities in more urban locations. This relatively frequent staff turnover requires local governments to periodically hire and train new staff. Another common factor found to slow ICM adoption in most countries was the interruption of financial and political support for ICM by election cycles and corresponding changes in administrative priorities. This has required PEMSEA staff to also periodically re-orient government officials and try to convince them to continue their predecessor's priorities to improve the environment.

The updated findings on ICM scaling-up and replication show that while broader adoption has taken place more quickly than we expected, environmental change at the regional level is still elusive after more than 25 years of support. This is partly because a larger geographical scale requires a corresponding larger temporal scale for significant change to be detected. At the same time, progress on ICM implementation has differed from case to case. ICM has not been fully adopted in many sites, and in many cases is resource-constrained, as mentioned above. In China where much time, effort and resources have been invested in ICM, improvements in water quality have been observed over the last decade (e.g. Zeng et al. 2019); it is reasonable to conclude that such significant investments have contributed to reducing environmental stress along the China's coast.

Key to measuring ICM's contributions is having systems in place for monitoring and reporting environmental change. Evidence of ICM's positive impacts has been reported at the local level, but some are anecdotal. Many sites lack the infrastructure and human capacities to monitor water quality. At the regional level, there is no overall system to track social and environmental trends related to ICM implementation. Despite several GEF-supported initiatives promoting monitoring and reporting systems in the region, these initiatives have been adopted partially or sporadically by participating governments. For example, in 2008, PEMSEA introduced the State of the Coast (SOC) Report, a self-reporting scorecard consisting of legal, institutional, economic and environmental indicators that have now been adapted for the municipal, provincial, national and regional levels. PEMSEA has been assisting ICM sites in putting together SOC reports, but many remain in various stages of development (Padayao 2018); only a few sites have published more than one report to track changes over time.

Broader adoption of local-scale interventions by stakeholders at different scales is necessary so that large-scale drivers are addressed. At the time of the evaluation in 2011, ICM was being gradually replicated from pilot sites to neighbouring municipalities or scaled up at the provincial level; the adoption of national regulations in support of ICM was in its early stages. Many of the countries that had gained broad endorsement of ICM from the executive branches of national government then are now in the process of getting ICM-related laws approved by the legislative branches. In some countries, implementing mechanisms, such as

intersectoral committees, are already being established. Key to this development is that most member countries have drawn on the experience of pilot ICM sites established in the first 10 to 15 years of GEF support to adopt legal, regulatory and institutional reforms that facilitate ICM adoption at a larger scale.

In China, lessons from the pilot sites have been leveraged through the China-PEMSEA Sustainable Coastal Management Cooperation Center (CPC), which provides training and support for ICM expansion in all coastal cities. Other countries such as Cambodia, Thailand and Indonesia are considering establishing similar centres. However, widespread geographical expansion of ICM does not automatically address large-scale drivers, many of which originate beyond coastal areas.

At least three ICM sites – also three of the longest-running ones supported by the GEF – have aimed to expand the geographical reach of the intervention to better match the ecological boundaries of the problem, which in this case is coastal pollution originating from both coastal and non-coastal areas.

One site is Xiamen in China where the local government linked with the existing Jiulong River Watershed Project (JRWP). Together they established a system to control pollution discharges in the river that supplied water to the city and discharged in Xiamen Bay. Using a payment-for-environmental-services (PES) arrangement, Xiamen transferred funds to the upstream municipalities in exchange for clean water. Ten years later in 2020, the PES system is still functioning and continues to result in improvements in the river's water quality. The second site is Manila Bay, where PEMSEA helped expand ICM in the cities surrounding the bay. The Manila Bay ICM initiative is currently linked with the country's River Basin Control Office to address pollution sources in the two major river basins of Pampanga River and Pasig River-Laguna de Bay through integrated basin management, stakeholder engagement and information management.

In the province of Chonburi in Thailand, the ICM programme expanded to include all 26 coastal municipalities from 2005 to 2009; by 2010, the 73 non-coastal municipalities had signed on to implement a Provincial Coastal Development Plan seeking to address upland pollution from pig and agricultural farms.¹⁰ The city of Sriracha, Thailand's ICM pilot site, provided training in non-coastal municipalities. Clusters of municipalities were to prepare ICM plans for consolidation into the provincial ICM plan. However, due to logistical constraints and the large number of local governments involved, only the 26 coastal municipalities developed ICM action plans; ICM practices, particularly in solid waste management, continued to be applied throughout the entire province, coordinated by the provincial government. An interagency and multi-sectoral Provincial Marine and Coastal Resources Management Committee was established in 2015, having jurisdiction over both coastal and non-coastal municipalities. An informal survey of the heads of 25 coastal and 21 non-coastal municipalities in 2015 showed that ICM for them created a "platform for intermunicipal dialogue and exchange of information" (Barnette and Wiwekwin 2018). We were, however, unable to reach local government officials to get an update on what has happened since.

While we do not have complete information on environmental outcomes in these three ICM sites, based on the developments post-evaluation, it appears that

it may indeed be possible to address large-scale drivers by scaling up a local-scale intervention such as ICM. In turn, implementation at higher scales has been made possible by the existence of enabling conditions at the corresponding higher scales, such as the regulations and intersectoral committees gradually established at the national and provincial levels after many years of learning from pilot sites.

From evaluating static long-term goals to evaluating development trajectories

The three main assumptions examined in this postscript remain mostly valid. However, the analytical model used in the evaluation had other misleading assumptions that have become evident in retrospection. First, in the introduction of the chapter, we stress the importance of assessing the progress made towards achieving impact. This assumes that “impact”, which is the goal of the intervention and assumed to take place sometime in the distant future, would remain unchanged. Second, the model adopted by the evaluation assumes that a successful intervention would put in place a set of conditions that, once achieved, would over time enable the achievement of the desired long-term goal.

However, interventions concerned with large complex systems such as the East Asian seas are better understood as efforts to steer such large systems towards a development trajectory that enhances the likelihood that desired benefits or services are generated. This trajectory is constantly influenced by multiple actors and factors, which introduces a high degree of uncertainty into the process. Over time, the predetermined conditions for success might no longer contribute to sustaining the desired benefits and services. Shifting circumstances and new information generated during implementation will require periodic reassessment of the conditions that are assumed to enable long-term change and, if necessary, a redefinition how “impact”—the long-term objective of the intervention—is defined. Since the system is inherently unpredictable, it is therefore more pragmatic to track its trajectory in providing benefits and services desired, rather than just the extent to which the pre-specified enabling conditions have been established.

The environmental problems of the South China Sea and the other East Asian seas are far from solved. Even after more than 25 years of GEF support, it may not be realistic to expect a dramatic improvement in the environmental health of these water bodies. Nevertheless, there is considerable anecdotal evidence that the types of adaptive approaches and models supported by the GEF help to steer the region towards a development trajectory that enhances the delivery of valued environmental and economic services.

PEMSEA’s long-term goals have not remained static. To ensure the desired level of environmental services, the overall regional goal was increased from 20% to 25% of the region’s coastline placed under ICM. Similarly, ICM practices were scaled up to address stresses originating from non-coastal municipalities. The conditions targeted for ICM adoption to enable long-term change have gradually been adapted to specific local and national circumstances, and are increasingly incorporating new approaches, sectors and stakeholders.

What we have yet to see is a better understanding of the extent to which the many local success stories translate into a broader development trajectory that will ensure the sustainability of the desired services at a regional scale in the East Asian seas. Systematic and regular State of the Coast reports at multiple scales could provide valuable information to assess the broader trajectory and make adjustments if necessary.

As we say in our conclusion almost ten years ago, the evaluator needs a sound understanding of the system; it is clear now that it is also important to understand how the system develops over time. However, it is not realistic to expect that we will ever have perfect knowledge to predict when, how and what long-term impacts will take place due to the unpredictable nature of complex systems. Often we may not see the outcomes we wanted or expected in the time that we predicted. After more than 25 years of GEF support, we expected to see some large-scale environmental changes in the region by now, yet there is no counterfactual region to allow us to say if the system's pace of change would have been faster or slower without this continuous long-term support. In this sense, trying to predict when long-term goals might finally be achieved is an almost futile exercise. More useful may be to periodically assess changes in the system and the extent to which interventions have adapted in ways that support the development trajectory towards desired benefits and services – which themselves could also change over time.

Notes

- 1 This sea is called the “South Sea”, 南海, Nánhǎi in China and the “East Sea”, Biển Đông, in Viet Nam. In September 2012, the president of the Philippines issued Administrative Order No. 29, renaming the “maritime areas on the western side of the Philippine archipelago as the West Philippine Sea”, to cover the areas within the Philippines’ Exclusive Economic Zone. “South China Sea” is the international name for this area as recognized by the International Hydrographic Organization and is therefore the name used throughout this evaluation.
- 2 These are the Convention on Biological Diversity, United Nations Framework Convention on Climate Change, the UN Convention to Combat Desertification, Stockholm Convention on Persistent Organic Pollutants and the Minamata Convention on Mercury. For more information on the GEF, see <https://www.thegef.org/about-us>.
- 3 The GEF strategy for the international waters focal area is largely focused on facilitating cooperative action among governments that derive benefits from a shared water body, such as a large marine ecosystem, lake, river or aquifer. The initial approach is the undertaking of a transboundary diagnostic analysis, through which threats to the shared water body that cross or extend beyond national jurisdictions (i.e. are transboundary) are identified based on available scientific data. The countries then jointly prioritize which of these environmental concerns to take action on. The priority environmental concerns and actions with which to address them are formalized through the countries’ agreement to a regional Strategic Action Programme.
- 4 Scale in this context is to be understood as the spatial, temporal, quantitative or analytical dimensions used to measure and study any phenomenon (Gibson et al. 2000).
- 5 A higher scale implies a greater geographical scope, and a greater number and types of interacting elements, such as ecosystems and social, political and economic sub-systems.

- 6 A list of large marine ecosystems similar to the South China Sea and Gulf of Thailand in terms of ecological, political and socio-economic characteristics was derived from Mahon et al. (2010).
- 7 These were the 1992 United Nations Conference on Environment and Development (the Rio Earth Summit) in Rio de Janeiro, the 2002 World Summit on Sustainable Development in Johannesburg and their precursor, the 1972 United Nations Conference on the Human Environment (Stockholm Sustainable Development Conference) in Stockholm.
- 8 The Impact Evaluation of GEF Support to International Waters in the South China Sea and Adjacent Areas was undertaken from 2010 to 2012.
- 9 Various figures have been published in 2018 and 2019 PEMSEA reports and proceedings of the project steering committee.
- 10 In the case of Chonburi, there was no major river draining into marine water bodies, only small streams and canals.

References

- Anderies, J. M., M. A. Janssen, and E. Ostrom, 'A framework to analyze the robustness of social-ecological systems from an institutional perspective', *Ecology and Society*, 9(1): 18–36, 2004.
- Babcock, R. C. et al., 'Decadal trends in marine reserves reveal differential rates of change in direct and indirect effects', *Proceedings of the National Academy of Sciences*, 107(43): 18256–18261, 2010.
- Barnette, P. and N. Wiwekwin, 'Integrated coastal management implementation and scaling up in chonburi province, Thailand, In: Chua, T.-E., L. M. Chou, G. Jacinto, S. A. Ross, and D. Bonga, eds., *Local Contributions to Global Sustainable Agenda: Case Studies in Integrated Coastal Management in the East Asian Seas Region*. Partnerships in Environmental Management for the Seas of East Asia (PEMSEA) and Coastal Management Center (CMC), Quezon City, Philippines, 469–482, 2018.
- Brondizio, E., E. Ostrom and O. R. Young, 'Connectivity and the governance of multilevel socioecological systems: The role of social capital', *Annual Review of Environmental Resources*, 34: 253–278, 2009.
- Cash, D. W. et al., 'Scale and cross-scale dynamics: Governance and information in a multilevel world', *Ecology and Society*, 11(2): 8–20, 2006.
- Davies, R., 'Scale, complexity and the representation of theories of change', *Evaluation*, 10(1): 101–121, 2004.
- Dulvy, N., R. P. Freckleton and N. V. C. Polunin, 'Coral reef cascades and the indirect effects of predator removal by exploitation', *Ecology Letters*, 7: 410–416, 2004.
- Folke, C., S. Carpenter, T. Elmqvist, L. Gunderson, C. S. Holling and B. Walker, 'Resilience and sustainable development: Building adaptive capacity in a world of transformations', *AMBIO: A Journal of Human Environment*, 31: 437–440, 2002.
- Folke, C., L. Pritchard Jr., F. Berkes, J. Colding and U. Svedin, 'The problem of fit between ecosystems and institutions: Ten years later', *Ecology and Society*, 12(1): 30–42, 1997.
- Folke, C. et al., 'Regime shifts, resilience, and biodiversity in ecosystem management', *Annual Review of Ecology, Evolution, and Systematics*, 35: 557–581, 2004.
- Gibson, C. C., E. Ostrom and T.-K. Ahn, 'The concept of scale and the human dimensions of global change: a survey', *Ecological Economics*, 32(2): 217–239, 2000.
- Gladwell, M., *The Tipping Point: How Little Things Can Make a Big Difference*, Boston, Little Brown, 2000.
- Global Environment Facility (GEF), 'Impact Evaluation of GEF Support to International Waters in the South China Sea and Adjacent Areas', Global Environment Facility Independent Evaluation Office (GEF IEO), Washington, DC, 2012.

- Gunderson, L.H. and C.S. Holling, eds., *Panarchy: Understanding Transformations in Human and Natural Systems*, Washington, DC, Island Press, 2002.
- Holling, C.S., 'Understanding the complexity of economic, ecological, and social systems', *Ecosystems*, 4: 390–405, 2001.
- Jackson, J.B.C. et al., 'Historical overfishing and the recent collapse of coastal ecosystems', *Science*, 293(5530): 629–638, 2001.
- Kinzig, A. P., P. Ryan, M. Etienne, H. Allison, T. Elmqvist and B.H. Walker, 'Resilience and regime shifts: assessing cascading effects', *Ecology and Society*, 11(2): 20–43, 2006.
- Liu, J. et al., 'Complexity of coupled human and natural systems', *Science*, 317: 1513–1516, 2007.
- Mahon, R., L. Fanning, P. McConney and R. Pollnac, 'Governance characteristics of large marine ecosystems', *Marine Policy*, 34(5): 919–927, 2010.
- Mayne, J. 'Contribution analysis: Addressing cause and effect, In: R. Schwartz, K. Forss, and M. Marra, eds., *Evaluating the Complex*, 53–97, New Brunswick, NJ, Transaction Publishers, 2011.
- Mee, L. D., H.T. Dublin and A. A. Eberhard, 'Evaluating the global environment facility: A goodwill gesture or a serious attempt to deliver global benefits?', *Global Environment Change*, 18(4): 800–810, 2006.
- Organisation for Economic Co-operation and Development (OECD), 'Glossary of key terms in evaluation and results based management, 2002, available at <http://www.oecd.org/development/peer-reviews/2754804.pdf>, accessed August 18 2013.
- Ostrom, E., 'A general framework for analyzing sustainability of social-ecological systems', *Science*, 325 (5939): 419–422, 2009.
- Ostrom, E., 'Beyond markets and states: Polycentric governance of complex economic systems', *American Economic Review*, 100: 1–33, 2010.
- Padayao, D. 'Enhancing accountability in environmental governance through the state of the coasts reporting system, In: Chua, T.-E., L.M. Chou, G. Jacinto, S.A. Ross, and D. Bonga, eds., *Local Contributions to Global Sustainable Agenda: Case Studies in Integrated Coastal Management in the East Asian Seas Region*. Partnerships in Environmental Management for the Seas of East Asia (PEMSEA) and Coastal Management Center (CMC), Quezon City, Philippines, 219–234, 2018.
- Pandolfi, J. M. et al., 'Are U.S. coral reefs on the slippery slope to slime?', *Science*, 307(5716): 1725–1726, 2005.
- Pawson, R., T. Greenhalgh, G. Harvey and K. Walshe, 'Realist review – A new method of systematic review designed for complex policy interventions', *Journal of Health Services Research and Policy*, 10(suppl 1): 21–34, 2005.
- Pawson, R., and N. Tilley, (1997). *Realistic Evaluation*, London, Sage, 1997.
- Sanderson, I., 'Evaluation in complex policy systems', *Evaluation*, 6: 433–454, 2000.
- Stern, E., N. Stame, J. Mayne, K. Forss, R. Davies and B. Befani, 'Broadening the Range of Designs and Methods for Impact Evaluations', London, Department for International Development, 2012, available at http://r4d.dfid.gov.uk/pdf/outputs/misc_infocomm/DFIDWorkingPaper38.pdf, accessed May 24 2013.
- Taleb, N. N., *The Black Swan: The Impact of the Highly Improbable*, Second Edition, New York, Random House, 2007.
- van den Berg, R. and D. Todd, 'The full road to impact: The experience of the Global Environment Facility fourth overall performance study', *Journal of Development Effectiveness*, 3(3): 389–413, 2011.
- Zeng, J., Yin, B., Wang, Y., and Huai, B. (2019). Significantly decreasing harmful algal blooms in China seas in the early 21st century. *Marine Pollution Bulletin*, 139, 270–274.