Clustering Based Assessment of Cost, Security and Environmental Tradeoffs with Possible Future Electricity Generation Portfolios

Yusak Tanoto[a,c], Navid Haghdadi[a], Anna Bruce[b] and Iain MacGill[a]
[a] School of Electrical Engineering and Telecommunications and Centre for Energy and Environmental Markets, University of New South Wales Sydney, Australia
[b] School of Photovoltaic and Renewable Energy Engineering and Centre for Energy and Environmental Markets, University of New South Wales Sydney, Australia
[c] Electrical Engineering Department, Petra Christian University, Indonesia

Abstract
The electricity sector has a key role to play in the sustainable energy transition. The falling costs of wind and solar PV have added to both the opportunities yet also challenges of balancing sometimes competing industry objectives of costs, security, and environmental impacts. This paper presents novel techniques for assessing possible future industry generation portfolios in three ways: 1) incorporating explicit metrics for energy trilemma objectives into modelling, 2) using the optimization process of evolutionary programming to map the solution space of 'high performing', near least-cost, portfolio solutions, and 3) applying boundary min-max cases and clustering to categorize these varied portfolios to better facilitate planning and policy making. We use an open-source evolutionary programming tool, National Electricity Market Optimiser, to assess possible future generation portfolios for Indonesia's Java-Bali interconnected power system. Our findings highlight the wide range of possible portfolios that might potentially deliver similar total industry costs, and their different security and environmental implications. In particular, additional solar photovoltaic deployment appears a low-risk opportunity to reduce costs and emissions compared to more fossil-fuel oriented mixes. Our novel techniques may be useful for the energy modelling community seeking to better understand and communicate complex, uncertain, and multi-dimensional choices for electricity industry planning.

Keywords: Clustering analysis; electricity industry; energy trilemma; generation portfolios; reliability; tradeoffs.

1. Introduction

The Paris Agreement in 2015 has created momentum for mitigation efforts from all jurisdictions, including developing countries, towards internationally agreed targets to avoid the worst impacts of climate change. The electricity industry has been recognized as a major contributor to climate change [1] and meeting these climate goals therefore requires transitioning electricity industries in many developing countries from heavily dominated by fossil-fuel based generation to low emissions generation future pathways. This transition is expected to include significant investment in large-scale renewable energy generation, with capacity expansion also required to meet a more dynamic growing electricity demand in industrializing countries. While the high cost of renewable energy technologies such as solar PV and wind has historically presented a barrier to a low emissions transition, the falling costs of these technologies now present a key opportunity, not just for environmental sustainability, but also for reducing industry costs and improving reliability by ensuring resource adequacy.

Electrical power systems can draw upon numerous potential generation options, including a range of low or zero operating emission technologies, and deliver almost all energy services, including those currently supplied by liquid fuels and natural gas. However, the physical characteristics of electricity, and the large interconnected networks that most economically deliver it, also present specific challenges for managing security of supply, particularly in the context of developing countries.

The high capital intensity and long lives of electricity sector assets, their differing contributions to system security, and varied environmental and wider societal impacts all pose considerable difficulties for energy planners seeking to balance social objectives of access and affordability, energy security and environmental sustainability. Indeed, the term energy trilemma [2] is often used to emphasize the potential tradeoffs involved across these objectives. A number of planning frameworks have used various quantitative metrics for each of these objectives to assess and compare the present sustainability of different electricity sector jurisdictions, changes in jurisdictions over time, and possible future scenarios Common industry metrics include total industry costs as a partial measure of affordability, expected unserved energy (USE) for security and carbon emissions (tCO_2) for environmental sustainability. These comparisons often highlight that only a few developed countries have established the highest-balance scores for all trilemma dimensions [3], [4]. For developing countries, managing the energy trilemma is generally even more difficult as electricity infrastructure is typically insufficient to meet growing demand within an acceptable range of reliability, budget limitations can reduce options (particularly those with high upfront costs) while many of these countries also face a scarcity of fossil fuel resources.

The extraordinary progress seen in wind and photovoltaic generation technologies over the past two decades has added to both the opportunities yet also challenges of jurisdictional industry planning to balance this energy trilemma. Costs have fallen to the extent that they are now competitive with conventional, largely fossil fuel-based options in some cases, while they have no, or at most very low, operating environmental impacts. However, they are both highly variable and only somewhat predictable, with no inherent energy storage [5], hence add to the challenges of maintaining security of supply [6].

A large and growing number of studies are undertaking generation capacity expansion planning with large scale renewable energy integration for different jurisdictions using various simulation and optimization methods and tools, which have been applied specifically to the task of assessing the total industry costs, achieved reliability and environmental outcomes [7], [8]. While linear programming, mixed integer programming, dynamic programming and evolutionary programming are widely utilized in those studies, the very large uncertainties associated with long term industry planning including future costs of different generation technologies, environmental policy drivers and possibly changing reliability preferences (particularly with new distributed energy technologies) are, if considered at all, typically addressed through the use of sensitivity studies and scenario analysis [9]. Other than that, those studies on the context of developing countries – which have very different levels of achieved reliability to that enjoyed by industrialized economies – have applied reliability targets for highly developed electricity sectors, for example by considering no unmet energy in typical studies of developing countries' long-term planning with bottom-up accounting model [9], which may inappropriately slant analysis outcomes towards higher cost solutions.

A recent study conducted evaluation of different energy scenarios that are supporting a more resilient low carbon energy system [10]. While energy trilemma index has been utilized in the analysis, this study was focused on the assessment of single optimum solution conducted using a mix-integer programming tool considering several generation scenarios and interconnection options. Moreover, key uncertainties surrounding future electricity industry planning, such as future costs and environmental policy drivers, however, have not been addressed. Other study has conducted assessment on the required CO_2 abatement costs from electrification and CO_2 mitigation tradeoffs for the future generation mix in Java-Bali grid system, Indonesia, given various emissions reduction targets [11]. While this study has stressed on the tradeoffs analysis between electrification and emissions reduction targets, none of the technical characteristics associated to variable renewable energy penetrations, such as temporal variability and / or uncertainty due to resource output fluctuation, is discussed and applied.

Few recent studies have explicitly incorporated uncertainties beyond the use of sensitivity and scenarios analysis within the context of generation expansion planning. Uncertainty regarding the costs of different generation technologies involving coal, OCGT and CCGT have been explored to obtain a range of expected generation costs of portfolios, expected CO_2 emissions, and standard deviation of cost using a Monte Carlo simulation platform, in which capital cost uncertainty, fuel price, carbon price, and demand uncertainty and elasticity have been taken into account [12]. Meanwhile, multiple uncertainties comprising electricity demand, capacity factor of variable renewables, and environmental policy have been considered in exploring the optimal mix of generation technologies using a two-stage stochastic model [13]. Other study considered capacity factor of variable renewables plus hydro to assess output probability distribution and generate renewable energy sources scenarios through Monte Carlo simulation combined with a deterministic generation expansion planning model to find optimal capacity mix [14]. While these studies have shown an in-depth and useful analysis in identifying and modelling uncertainties around electricity industry future, however, these studies focused on the probability density approach and frontier optimal generation mixes.

Based on this literature review, our study seeks to fill these following three major gaps and limitations – as key concerns with much of the work undertaken to date around electricity industry planning with high variable renewable penetrations, particularly in the context of developing countries. These three gaps are, firstly, none of these studies explicitly and comprehensively explore the tradeoffs, yet also potential synergies, of future generation portfolios across the entire energy trilemma objectives of cost, security and environmental impacts. This is despite such tradeoffs being a key challenge for policy makers and planners, and despite the considerable variation in prioritization that may be seen across jurisdictions. Secondly, many of existing studies have used optimization techniques that can be problematic for exploring the implications of future uncertainties for energy trilemma objectives, as optimization techniques typically seek a single 'lowest' cost solution. However, this lowest cost solution will depend on the cost and constraint assumption 'inputs'. A relatively small change in estimated future costs for particular generation technologies (PV is a particularly pertinent example here) may entirely change the 'least cost' outcome. And it may be that a very different generation portfolio is only slightly more expensive but has extremely desirable characteristics for planners in terms of, for example, local industry participation and reduced reliance on fuel imports. Thirdly, none of the existing studies have utilized clustering techniques to present a highly complex choices as a result of managing energy trilemma objectives in the area of electricity generation expansion planning. Where efforts to map such tradeoffs have been undertaken partially in very few studies, there is the challenge of presenting potentially highly complex choices in a way that is useful to planners and energy policy makers.

Our study therefore presents a novel approach that allows a comprehensive assessment and exploration of energy trilemma key metrics performance of electricity generation planning with high variable renewable penetrations using three integrated methods. Firstly, our modelling approach explicitly explores energy trilemma tradeoffs between estimated future costs, reliability and environmental impact rather than treating one or more of them as constraints. Secondly, we use an optimization tool based on evolutionary programming that can, with modest modifications, provide far more information on the shape of the solution space by explicitly mapping 'near optimal' generation mix solutions. Other optimization techniques such as linear programming and dynamic programming do provide some guidance on the near solution space – for example, shadow pricing binding constraints, and mapping all state transition costs. However, evolutionary programming techniques explicitly solve the costs of a wide ‘population’ range of possible generation solutions as they evolve better solutions, and solutions that are only slightly more expensive than the least cost solution can be analyzed through ‘cost relaxation’ as we explain below. Thirdly, we apply a range of techniques including clustering analysis to present this highly complex ‘near optimal’ generation mix solution space in more informative ways for electricity planners and policy makers.

To demonstrate our proposed approach, we apply it to the particular context of the Indonesian Java-Bali interconnected power system. The Indonesian electricity industry presents an interesting and important opportunity for analysis. The world’s fourth most populous nation is in a period of rapid economic growth that is stressing the existing industry and demands major capacity expansion in the coming decades. While Indonesia has plentiful coal, it also has excellent renewable, particularly hydro, geothermal and solar, resources. Reliability is a challenge while the environmental impacts of its coal dominated generation sector are also an issue, particularly if future developments focus on coal generation expansion. The government has developed a power plan that incorporates growing renewables yet a continued reliance on coal [15]. Whilst other groups have presented alternative visions for Indonesia’s electricity industry future [16], our study is intended to make a further contribution to the deliberations of Indonesian electricity industry planners and policy makers in an increasingly uncertain context of planning.

Our study makes three new contributions to the body of knowledge around electricity industry planning, particularly to the energy modelling community’s suite of methods and tools for exploring possible sustainable electricity industry futures as we: 1) explicitly map the near-optimal solution space of generation mixes that perform nearly as well as the ‘optimal’ solution – and which may of course actually could be better given future uncertainties; 2) explicitly categorize these high performing generation portfolios according to some key metrics for the energy trilemma; and 3) use clustering to better understand the key performance tradeoffs across these metrics. In summary, the innovation of our study is in the use of cost relaxation to get many reasonably low-cost solutions and then clustering techniques to identify key aspects of this range of solutions. We also hope that our study can provide some specific insights for Indonesian energy planners and policy makers as well as wider communities contemplating their options for a more sustainable energy future.

The rest of this paper is organized as follows. We outline our methods in Section 2 while Section 3 presents results obtained from the case study of Indonesia’s Java-Bali grid. Finally, concluding remarks and thoughts for possible future work are presented in Section 4.

2. Methods

2.1. The methodological framework

In this study, we develop a framework to assess possible future electricity generation portfolios given the multiple objectives of the energy trilemma, and significant uncertainties, particularly related to future technology costs. The framework integrates two different methods. To incorporate uncertainty, multiple near least cost generation portfolios are collected from a stochastic optimization model. To assess performance across multiple objectives, clustering is used to examine the performance of the generation portfolios with respect to energy trilemma key metrics. The conceptual framework is presented in Fig. 1, while the framework’s components are further elaborated in subsequent sections.

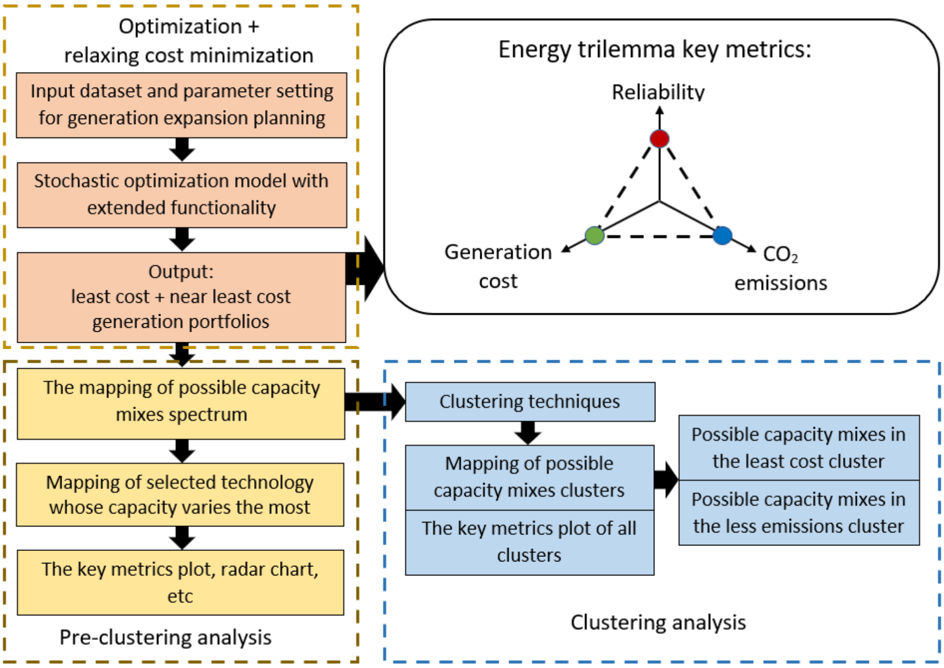


Fig. 1. The framework for assessing possible future generation portfolios.

2.2. Simulation overview and dataset

Our method for exploring possible future electricity industries involves first selecting a set of possible generation options for consideration. Some of these options are context specific – for example, our Indonesian Java-Bali grid case study incorporates coal, geothermal, power constrained hydro, coal-fired plant, Combined Cycle Gas Turbine (CCGT), Open Cycle Gas Turbine (OCGT), biomass, solar PV fixed axis and onshore wind farm options. The Indonesian coal-fired generation mix is coal-dominated. However, all of these generation options, including some small-scale grid-connected solar PV have been deployed in Indonesia, and later on followed by a breakthrough project of 75 MW wind farm located in Sidrap, Sulawesi island [15], and the latest commissioned of 60 MW wind farm located in the same island, while a number of assessments have highlighted some reasonable and promising solar and wind generation options, as mentioned in Indonesia's renewable energy prospect report [5], Indonesian wind map [17], and the catalogue for Indonesian power technology dataset [18]. Estimated annualized capital costs, fixed O&M, fuel cost, variable O&M, and the carbon emission intensity of each option is specified. Furthermore, any variable renewables require location specific generation profiles (typically 30-60 minutes interval over at least a year). An estimated demand profile of equivalent time steps and duration is also required.

Evolutionary programming tools generally create a feasible population of possible mixes of generation technologies and locations and then simulate their operation over the year or more of dispatch. Reliability may be set by putting a high price on any unserved energy, or as a constraint for feasible solutions. Similarly, carbon emissions can either be priced ($/t$CO_2$) or set as a constraint for the overall power system over the year. Evolutionary mechanisms are used to evolve this population of feasible solutions towards a least cost generation mix. Importantly, this evolving mix of generation portfolios can be tracked.

In this study, we use an open-source evolutionary programming-based techno-economic optimization model, National Electricity Market Optimiser (NEMO), as the primary simulation tool. Written in Python[1], the source code can be found in [20]. Like the majority of capacity expansion models, NEMO is designed to search for a single least cost generation investment option to satisfy the constraints applied. NEMO has been used in this way to model high renewables scenarios for the Australian National Electricity Market [19] and Indonesia [21]. However given the uncertainties associated with future costs and opportunities, particularly estimates of technology costs, it is important for planners to appreciate the characteristics of a set of portfolios that can potentially achieve the planning goals at near least cost, rather than a single modelled solution. This study, therefore, uses a novel approach, extending the tool functionality allow assessment of a large number of near least-cost generation portfolios with total costs that fall within 5% of the least cost solution found by the model. The study also adapts the tool to the generation expansion planning context of Indonesia's Java-Bali 2030 electricity grid.

NEMO contains a chronological dispatch model that is used to test portfolios of conventional and renewable electricity generation technologies and uses an evolutionary programming approach, Distributed Evolutionary Algorithms in Python (DEAP), to search for a near-least cost solution. DEAP is an evolutionary computation framework implemented in Python covering most common evolutionary computation techniques such as genetic algorithm, particle swarm optimization, and differential evolution [20]. NEMO implements the Covariance Matrix Adaptation Evolution Strategy (CMA-ES), a stochastic method for solving continuous domain optimization of non-linear non-convex function problems [22]. Fig. 2 shows the generic optimization framework of NEMO, which is unveiled by exploring the source code in [20], and the extension conducted in this study to the original framework as we enhance the functionality of the tool. It is intended to provide readers with an idea on the foundation of our work and our new contribution to the field area.

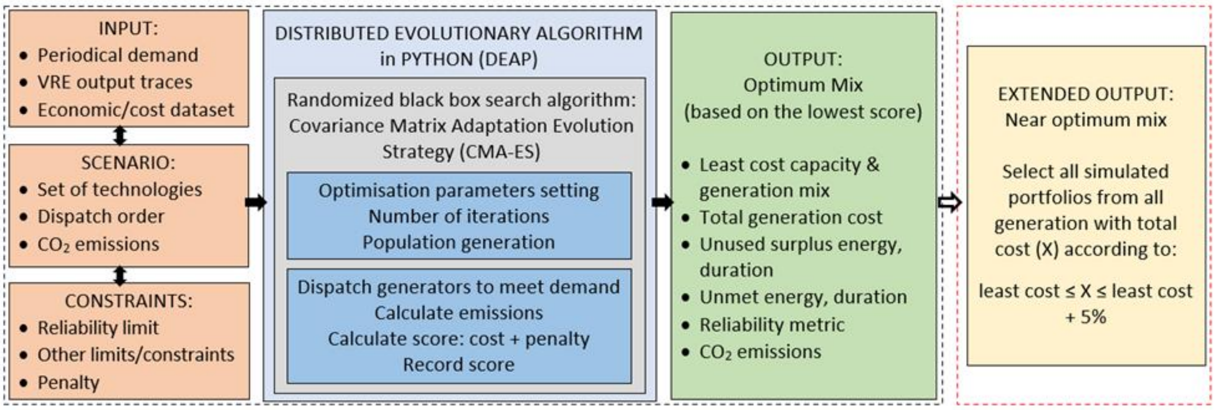


Fig. 2. The generic optimization framework in NEMO [20] and the extension in this study.

The dispatch program in NEMO takes inputs such as the projected hourly demand, hourly solar and wind output traces, coal and gas emissions intensity, discount rate, fuel and technology costs data and dispatches all generation technology candidates according to the predetermined dispatch order to meet the projected demand in each particular hour, applying a number of constraints and limits, including the predetermined reliability standard, technology capacity built limit, hydro energy generation limit, and

[1] Python is an open-source high-level programming language which is interpreted, interactive, object-oriented. It has a large and comprehensive standard library to support the development of a wide range of applications. Python is community driven and has an organization called The Python Software Foundation (PSF).

other policy-based constraints. NEMO searches for the least cost generation portfolios using the CMA evolution strategy, for a specified maximum number of iterations. In each generation of the search, several populations are established by the algorithm, and total generation cost (operational and capital) along with penalty (if any), any unused surplus energy, achieved unserved energy, and total CO_2 emissions are calculated for each population within each generation. The algorithm then compares the least cost solutions obtained in the current generation with the previous, to find the all-time least cost. In this study, NEMO is adapted to produce set of near least cost solutions by relaxing the cost constraint such that all generation portfolios with total generation cost fall within 5% least cost solution are retained.

Given the low penetration of variable renewables in the Java-Bali grid, our study does not consider grid capacity to manage variable renewable energy, the impact of distributed energy or energy storage technologies. This is aligned with, and allows the results to be compared with the published planning documents for the Java-Bali electricity grid. Nevertheless, technical integration studies would be beneficial to understand the characteristics, potential and limitations of the modelled scenarios, while exploration of the role of distributed energy and energy storage technologies would certainly add value in future studies.

We simulate 2030 generation portfolios for Indonesia's Java-Bali electricity grid, using 2015 hourly grid demand as a baseline. Data for gridded hourly PV and wind power output traces across the Java-Bali region are obtained from Renewables Ninja, an online renewable energy simulation tool [24] and Indonesia wind prospecting map [17], respectively. We select six locations dispersed within six provinces in the Java-Bali region to deploy solar PV and wind by considering the temporal and spatial variability, capacity factor, and potential output of both technologies [21]. We apply capacity-build limits for geothermal and hydro in 2030 according to [5]. NEMO's two key parameters for evolutionary optimization, i.e. number of generations and initial standard deviation of the distribution, are set at 100 and 2, respectively, as in [19]. We prioritize dispatch of variable renewable energy generation with the lowest variable O&M cost, then determine the hourly dispatch order of the synchronous generation candidates according to operating cost: geothermal, followed by hydro, coal, OCGT, CCGT, and biomass. For technology cost dataset, we use 2030 mid-level (base case) technology costs which are gathered from official sources [18], [23] as presented in Table 1. These are reasonably comparable to other international databases.

TABLE I
The 2030 Indonesian Java-Bali grid mid-level generation technology cost components

Technology	Capital ($/kW)	Fixed O&M ($/kW-year)	Variable O&M ($/MWh)
Geothermal (G)	3,200	16.7	0.7
Hydropower (H)	2,000	35.8	3.8
Coal fired (C)	1,360	35.8	3.8
OCGT (OC)	400	22.5	3.8
CCGT (CC)	710	22.5	3.8
Biomass (B)	1,600	43.8	6.5
Solar PV fixed axis (PV)	610	12.5	0.4
Onshore wind (W)	1,310	52	0.8

2.3. Modelling scenarios and assumptions

In this study, we firstly use a fixed USE limit of 0.005% as a reliability constraint, which is arguably a high reliability standard for the context of a developing country electricity industry. Secondly, we apply carbon prices of \$0/tCO_2 (CP0), \$30/tCO_2 (CP35) and \$60/tCO_2 (CP60) to reflect possible future policy settings to achieve Indonesia's emissions reduction commitments, but also to capture some aspects of uncertainty around future coal and other fuel costs, and costs associated with financing these increasingly risky projects. It should be noted that there is no specific carbon policy for the Java-Bali grid, however there is a government target of 23% renewable energy share on generation mix by 2025 countrywide [25]. We do not consider transmission network investment requirements.

Assumptions incorporate in our generation planning model include a non-synchronous penetration of 0.75, no minimum capacity reserves constraint, a 5% discount rate on annualized technology capital cost and a conservative annual demand growth of 5% from 2015 to 2030 [21], which is based on the historical average annual growth during a decade before and several years after 2015. As consequences, we obtain an energy demand profile of 346.5 TWh and 50 GW peak load for our modelling. Meanwhile, coal and gas prices in 2030 are assumed to be \$3.5/GJ and \$10.9/GJ, respectively [5], [6].

2.4. Clustering analysis and evaluation of clusters

Clustering analysis has been increasingly applied in some areas of electricity industry, such as in studies around load profile identification and demand estimation incorporating different datasets, such as load time series [26], smart metering data [27], demand profiles [28], occupant activity data [29], in addition to other applications covering smart energy system [30] and air-conditioning energy performance [31]. However, literature in the topic of clustering application for electricity generation expansion planning, is currently lacking. A recent clustering study in the area of capacity expansion planning is related to capturing the effect of variable renewables and energy storage [32] and time-period clustering for optimal capacity expansion planning with storage [33].

Clustering techniques are applied in this study to map various possible solutions lead to secure-affordable-low emissions technology mixes. This study primarily applies k-means clustering technique considering to its simplicity, efficiency, expandability, and ability to handle big data despite few possible drawbacks shown in some cases, such as sensitivity to the initial selection of cluster centers and chance to produce local optimum solution as mentioned in the study using smart meter data [27], demand profiles [28], and study on residential air conditioner control [34]. The step-by-step working principle of k-means clustering, in which the solution space is divided or partitioned into several cells containing data points with similar patterns, is summarized in [35]. Neural-network based-Self Organizing Map (SOM) is then used to compare the clustering results. The overall SOM algorithm is summarized in [36]. Neural network based-SOM clustering has been applied, for example, in estimating load for microgrid planning [73] and calculating load profile [38].

We apply the Calinski-Harabasz (CH) [39] and Davies-Bouldin (DB) [40] rules to evaluate the optimal number of clusters. In principle, these methods work by applying two important criteria, cluster compactness and separation. The first criteria measure the closeness of the member in each cluster to reflect cluster compactness while the second criteria measure the distance between clusters to reflect separation among clusters. The optimum number of clusters according to CH rule is detailed in [41]. DB rule measures the average of similarity between each cluster and its most similar one [42]. To follow the evaluation criteria, better number of clusters is represented by the lower-resulted DB index. Although there are some other evaluation rules available, the decision regarding number of clusters is eventually rather subjective depending upon the objective of analysis and details of the variability that the users would expect.

2.5. Adoption of capacity mix parameters into energy trilemma

For each generation portfolio, the simulation of a year of operation provides three representatives 'energy trilemma' metrics – achieved reliability (% USE), total industry costs - both capital and operating - ($/year) and carbon emissions (tCO_2). We use CO_2 emissions as a representative environmental metric since CO2 emissions represent the most significant long-term environmental risk. While coal-fired power plants, for example, release other greenhouse gasses such as methane and nitrous oxide, many official electricity generation planning and statistics, reports and other literature use CO2 emissions as an immediate representation of the air pollutants released by electricity generation and its value in quantifying the impact to environment. These are not complete representations of the trilemma; industry costs in particular do not capture the complexities of equity and affordability [1] while there are environmental impacts other than carbon emissions associated with different technology choices.

Although we set a USE constraint of 0.005%, this is an upper constraint and it is of course possible that candidate generation portfolios deliver lower USE. The conversion of the numerical values from actual achieved USE into a security score (magnitude) is, for simplicity, a linear interpolation from USE 0.000 to 0.005 equivalent to security scores from 100 to 75. We only consider security scores between 75-100 in this modelling exercise given the relatively high constraint. We conduct 2 steps of values conversion through normalization and scaling procedure to obtain the score for economic and environmental sustainability dimensions as follows:

(i) Sort candidate generation portfolios from the smallest to the largest value.
(ii) Determine the position (magnitude) of each value in the trilemma by normalizing the sorted values as following equation:

$$Normalized\ value_i = 1 - \frac{Old\ value_i}{Old\ value_{max}} \quad (1)$$

(iii) Determine the scale of the normalized values. The minimum and maximum scale for economic and environmental sustainability dimensions are set to 25 and 100, respectively, for visual clarity in the presentation of results. This provides the scale range of 75. The scaling operation to obtain the new values within a certain range of old values is carried out using the following equation:

$$New\ value_i = \left[\frac{(Old\ value_i - Old\ value_{min}) \times New\ range}{Old\ range}\right] + Scale_{min} \quad (2)$$

$$Old\ range = Normalized\ value_{max} \quad (3)$$

$$New\ range = Scale_{max} - Scale_{min} \quad (4)$$

There is one further complexity in industry costs – the treatment of carbon costs. The role of carbon pricing is to change the relative competitiveness of different generation technologies by changing their operating costs according to their emissions intensity. While it is of course possible to then include these costs (price times total annual emissions) in total industry costs, these costs do actually represent carbon revenue (CR) that can go towards helping pay for other industry costs or reduce taxes elsewhere. We therefore distinguish between industry costs in some of our analysis. Analyses in this study are conducted according to the following steps:

(i) Relaxing cost minimization: Using NEMO, we determine the set of candidate generation portfolios who costs fall within 5% of the least cost solution for a predetermined 0.005% USE limit and each carbon price scenario.
(ii) Mapping candidate generation portfolios.
(iii) Pre-processing (pre-clustering analysis 1): We analyze the candidate portfolios that have the highest and lowest PV capacity altogether with the mixes with the highest and lowest coal capacity, the highest and lowest gas capacity and the least cost mix. The three associated parameters of the mixes, i.e. cost, USE and CO_2 emissions are then mapped and analyzed as well. In this regard, we also analyze the realized costs with CR component included and without CR component included.
(iv) Pre-processing (pre-clustering analysis 2): We convert all associated parameters of the mixes (cost, reliability, emissions) into 3 dimensions of energy trilemma and map all values of the three dimensions of all seven portfolios (the highest and lowest PV, coal, gas and least cost portfolios) using radar charts. The charts depict the characteristic of different generation portfolios with respect to all dimensions in energy trilemma.
(v) Clustering analysis 1: We apply the k-means clustering technique into the 5% least cost technology mixes with different CPs to obtain group of technology mixes. We consider 6 clusters and use the k-means clustering algorithm in MATLAB. We use 1,000 repetitions to obtain a stable clustering membership number and configuration for each carbon price scenario. For comparison, we also use SOM clustering technique using nctool in MATLAB. As we obtain very similar results using SOM, we only present the results obtained from k-means clustering. Applying the CH and DB rules (briefly discussed earlier in section 2.4) to evaluate the number of clusters to be used, the suggested number of clusters based on these rules are mostly fall between 4 to 6 clusters. In this study, we choose 6 clusters to better capture the variability of the associated parameters that possibly emerge from the analyses.
(vi) Clustering analysis 2: For each CP, we analyze the 3 associated parameters of each cluster using the boxplot and compare cost and CO_2 emissions, as the two key parameters (given the reliability level is on the expected range) of all clusters. From this point, 2 distinct clusters can be identified, i.e. the cluster with least cost mix and the cluster with less emissions mixes.

3. Results and discussions

3.1. Base case results – least cost capacity and generation mix

We first solve least cost generation portfolios for the three different carbon price scenarios and the 0.005% USE limit. We save all feasible generation portfolios during the search whose overall costs ($/year) fall within 5% of the eventual least cost solution. Note that NEMO can incorporate existing generation but undertakes what is effectively a single investment step – in our case for 2030, rather than a sequence of investment steps.

Results for the 'least' cost generation portfolio with a 0.005% reliability constraint are shown in Fig. 3 for the three carbon price scenarios. Note that there are constraints on maximum hydro and geothermal capacity, as proposed in [21]. For CP0, there is substantial PV generation capacity in the least cost mix, but coal still predominates, with only a limited role for gas. CP35 sees more than double the PV but little other change in capacity. For CP60, PV capacity barely changes but wind generation now enters the mix, as does CCGT. Renewables climbs to over 70% of generation. These results might suggest certain choices and strategies for policy makers. However, the question arises of how sensitive they are to the specific 2030 cost assumptions, and what other generation mixes might offer relatively similar industry costs but perhaps quite different policy pathways.

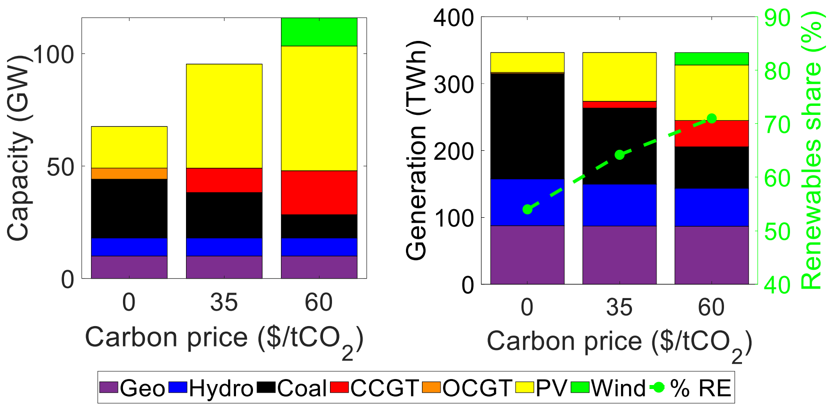


Fig. 3. (left) Least cost capacity mix and (right) least cost generation mix with fixed 0.005% USE limit for all CPs.

Results obtained for the case with no carbon price (CP0) are intended to reflect the latest single-scenario planning made by PLN [25], in which no carbon price is considered, to provide a basis for comparison and as a baseline for policy development. We therefore incorporate similar generation technology candidates to the published planning document. The results show that the renewables generation share by 2030 could achieve more than Indonesia's 23% renewables target [25], with solar capacity within the technical potential for Indonesia [5], [43]. This is in contrast to official planning, which shows only 18% share of renewable energy, while our least cost portfolios also show coal capacity less than that in the planning document (Fig. 3).

By achieving about 55% renewable energy share in the generation mix, notably including a high solar PV penetration as shown in Fig. 3 (right), our least cost mix results in 146 $mtCO_2$, while the CO_2 emissions of the government's BAU projection for 2030 (for the Java-Bali region) would be around 397 $mtCO_2$[2]. The targeted 29% reduction commitment would correspond to approximately 282 $mtCO_2$. Our scenario results in approximately 48% less emissions.

This highlights concern regarding Indonesia's pathways to de-carbonize the electricity industry in the future [44], and the role of coal. According to Indonesian government regulation no. 79/2014 on National Energy Policy, the coal share in the primary energy supply mix should be minimum 30% in 2025 and minimum 25% in 2050 [45]. However, based on the results presented in our study, we see a promising opportunity to reduce emissions at low cost.

[2] This value is an approximate amount of CO_2 emissions in Java-Bali electricity grid in 2030. We calculate this value based on the energy related emissions of 1,669 $mtCO_2$, which is specified in the Indonesian Nationally Determined Contribution to United Nations Framework Convention on Climate Change (UNFCCC). Reducing the energy related emissions by about 15%, we get 1,418 $mtCO_2$ to exclude non CO_2 greenhouse gas emissions, then reducing this CO_2 related emissions by 60%, and finally by another 30%, we get the approximate CO_2 emissions for electricity generation sector in Indonesia and in Java-Bali grid, respectively.

3.2. Fixed USE limit at 0.005%, cost relaxation up to 5% least cost technology mix

As we obtained a range of different least cost technology mixes presented in Fig. 3 (left), we expand each one of these three mixes into a spectrum of near least cost generation portfolios as shown in Fig. 4. This spectrum of portfolios, for each carbon price, is obtained by relaxing the overall industry cost up to 5% of that of the least cost mix presented in Fig. 3 and collating all generation portfolios that fall within this cost range.

The spectrum of near least cost portfolios, for example at CP0, consists of some 600 technology mixes. We obtain a wide capacity range across coal OCGT, solar and in some cases wind capacity. While geothermal and hydro capacity in 2030 remain at maximum potential for all carbon prices up to 10 GW and 8 GW, indicated with a downside arrow, the simulations result in both smaller and greater capacities for other technologies, indicated with two-way arrows.

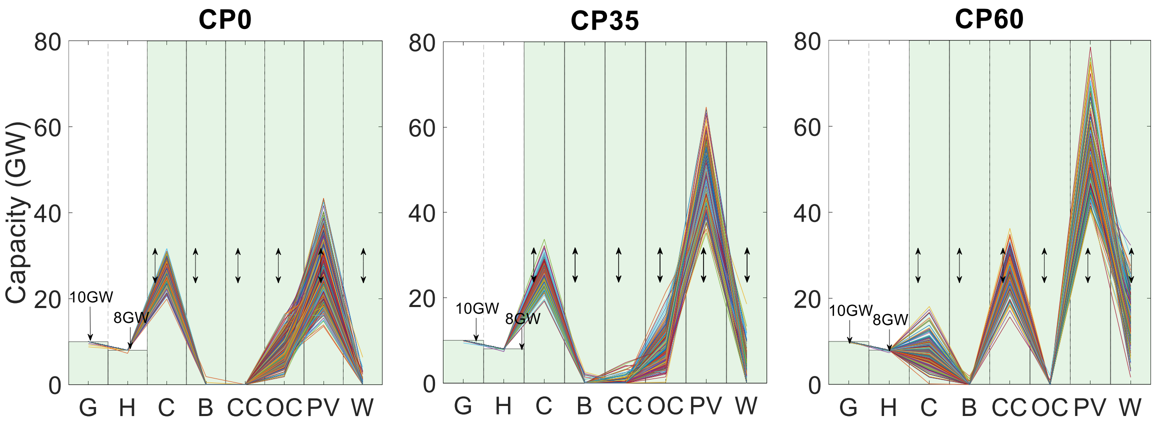


Fig. 4. Possible technology mix by relaxing cost up to 5% least cost mix and by fixing USE limit at 0.005% for CP0 (left), CP35 (middle) and CP60 (right).

The spectrum of technology mixes highlights the potentially wide range of generation investment futures. At CP0, there are evident options for greater or lesser coal, more OCGT and much more or less PV. At CP35, the role of gas including now CCGT could be quite substantial and the range of PV capacity increases and its range narrows. At CP60, coal capacity falls, CCGT climbs, PV capacity climbs further and wind becomes an option (from negligible to over 30 GW) across all candidate generation mixes. It is clear that the spectrum provides a richer set of insights for policy makers regarding their options and allows wider policy considerations to come into play – for example, concerns about future gas availability, or social acceptance of wind generation. However, it is less clear how these choices map to each other given that there are numerous candidate generation mixes presented, and choices in one technology capacity will, to at least some extent, dictate capacity choices in others.

3.3. Pre-clustering analysis

To better understand the tradeoffs, Selected candidate generation mixes, characterized according to those technologies whose capacity varies the most, are shown in Fig. 5. For CP0, it is interesting to note that there is not a very strong tradeoff between coal and PV capacity; indeed, the highest coal capacity mix actually has slightly more PV. The reverse also holds across the highest and lowest capacity of PV, which do not see great variation in coal capacity. The clearest tradeoff would actually appear to be OCGT and PV. For CP60, by comparison, there is an evident tradeoff between coal and CCGT generation, as well as with PV generation.

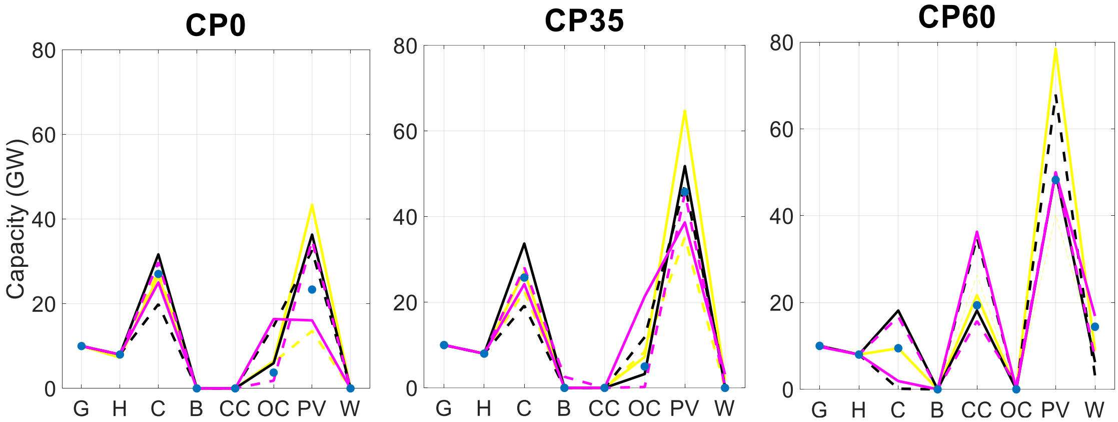
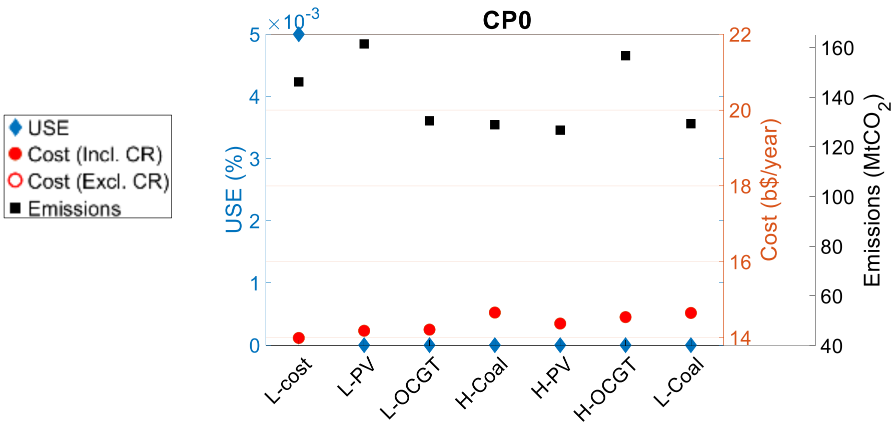


Fig. 5. Generation mixes with the highest and lowest possible capacity of PV, coal and gas, as well as the least cost mix, for each CP scenario.

These highest and lowest generation mixes for each key technology with 5% of least cost are further investigated in terms of emissions and USE in Fig. 6. For CP0, it is notable that the lowest cost mix has the highest USE, and higher emissions that all except the L-PV and H-OCGT cases. It appears that for around \$400m/year it would be possible to reduce emissions by around 20mtCO_2/year by deploying more PV. For CP35, the H-PV mix has mid-range costs yet the lowest emissions and USE. For CP60, the least cost mix is considerably lower cost than the other options and has mid-range emissions (they are of course priced at \$60/tCO_2) but lower reliability than all the other options. Excluding the carbon price revenue from total industry costs, on the basis that this money can be used to compensate energy consumers, has interesting implications in making the higher emission H-coal mix more attractive. When carbon revenue is excluded from industry costs, it is also interesting to note that CP35 makes the least cost mix only around \$200m/year more expensive, yet reduces emissions by over 25mtCO_2/year, which would seem to represent relatively low-cost abatement. CP60 costs around \$2.4b/year but reduces emissions by around 75mtCO_2/year, considerably higher cost abatement.



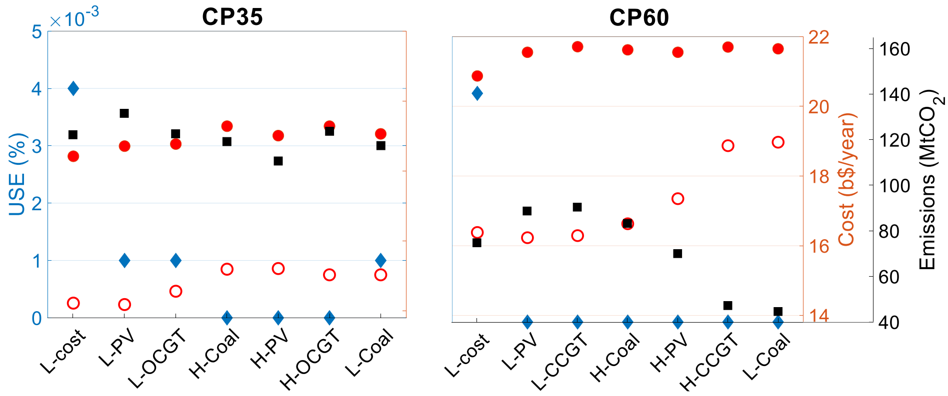


Fig. 6. The mapping of USE, total cost, and CO_2 emissions of generation mixes with the highest and lowest capacity of PV, coal, and gas, as well as the least cost mix for the three CP scenarios.

These outcomes for the min-max technology mixes can also be characterized using the trilemma indicators outlined earlier, and then plotted with radar diagrams as presented in Fig. 7. Note that these plots use industry costs excluding carbon price revenue, as this could be argued to involve double counting of the environmental impacts. Also, as noted above, the security metric for all these candidate generation mixes only varies between 75 and 100 reflecting the very low USE seen in all cases.

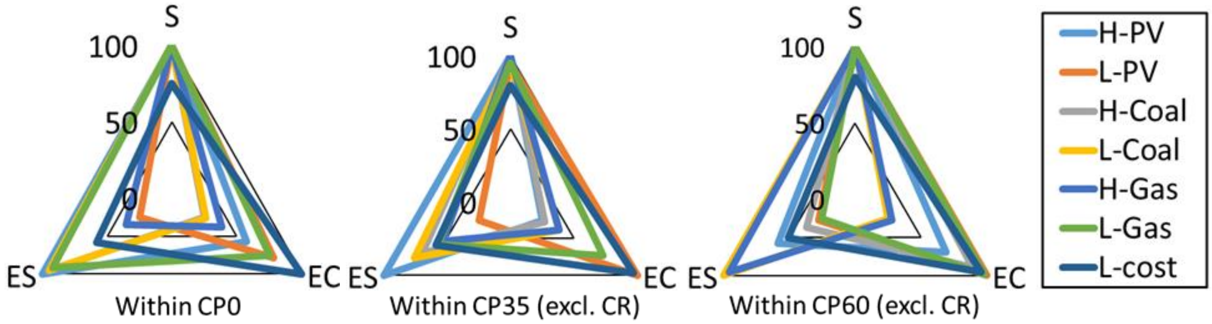


Fig. 7. Radar charts of security (S), economic (EC) and environmental sustainability (ES) dimensions for each min-max technology (economic indicator is calculated excluding CP revenues from industry costs).

These radar diagrams nicely illustrate the tradeoffs involved between choosing particular min-max technology mixes. There are no clearly superior candidate generation mixes across all the trilemma dimensions. However, there are a number of candidate mixes which would seem to have secure-affordable-low emissions alternatives – for example, the L-OCGT mix in CP0 and CP35 and the L-cost in CP60.

3.4. Results on clustering analysis

Considering the min-max technology deployment mixes is useful in bracketing the possible variation in mixes whose costs fall within 5% of the least-cost mix. However, to better characterize the solution space we use k-means clustering to group all solutions into six clusters. Figure 8 (left) shows 6 capacity clusters for CP0 case. The y-axis scale limits are made equal for all clusters to enable quick visual comparison of clustering patterns. The mean capacity mix for each cluster is shown using a black color line to help identify the spread of each of the clustering patterns.

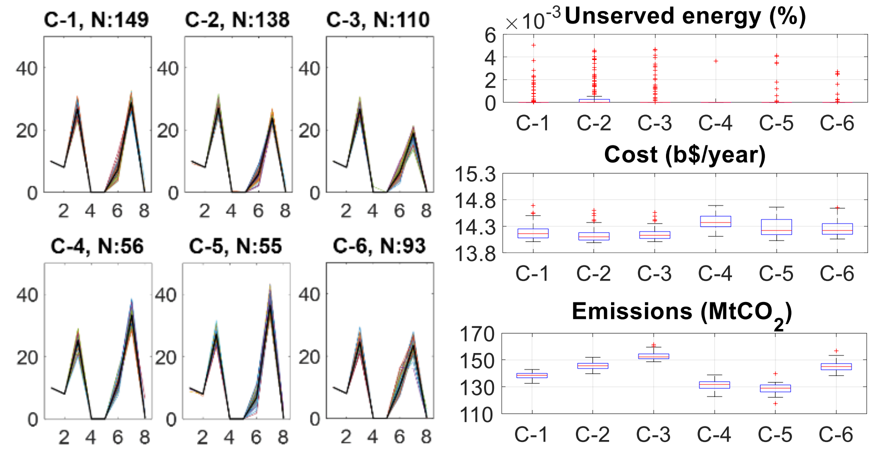


Fig. 8. (left) Six capacity clusters (y-axis = GW of capacity for each available generation option along the x-axis) and (right) unserved energy, cost and CO_2 emissions of all clusters CP0.

As shown in Fig. 8 (left), all clusters are characterized with little variations in coal capacity yet considerably more on PV. Higher capacity mixes of PV are clustered together in cluster 5 with a wide range of OCGT. This differentiates cluster 5 from cluster 4 despite both clusters having a similar pattern as seen from their mean capacity. Capacity clusters are further identified based on their USE, cost and CO_2 emissions, and the range of these are shown, and can be compared, in Fig. 8 (right). Cluster 2 has the lowest average cost. However, cluster 1, which has a similar range of costs to cluster 2, has a lower range of CO_2 emissions. The range of generation capacities in these two clusters are shown Fig. 9. It is interesting to note that a modest increase in costs allows greater PV and OCGT deployment.

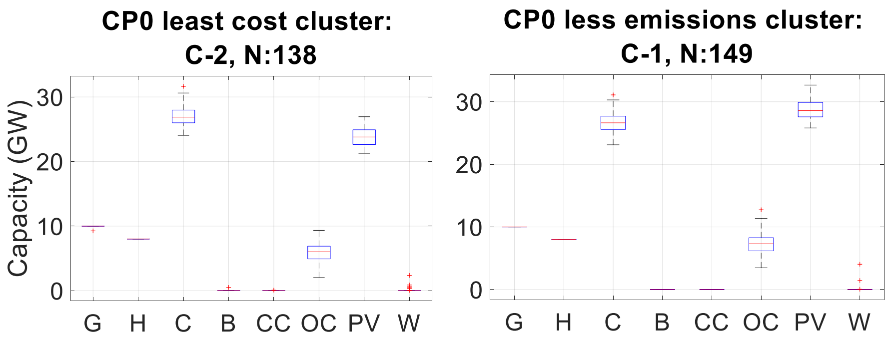


Fig. 9. (left) Generation technology mixes in the least cost cluster (C-2) and 'less emissions' cluster (C-1) for 5% least cost capacity mix with CP0.

Using the same analysis approach, clustering results of the technology mixes in CP35 and CP60 along with the results around least cost mix cluster and less emissions cluster identification are presented in Fig. 10 to Fig. 13.

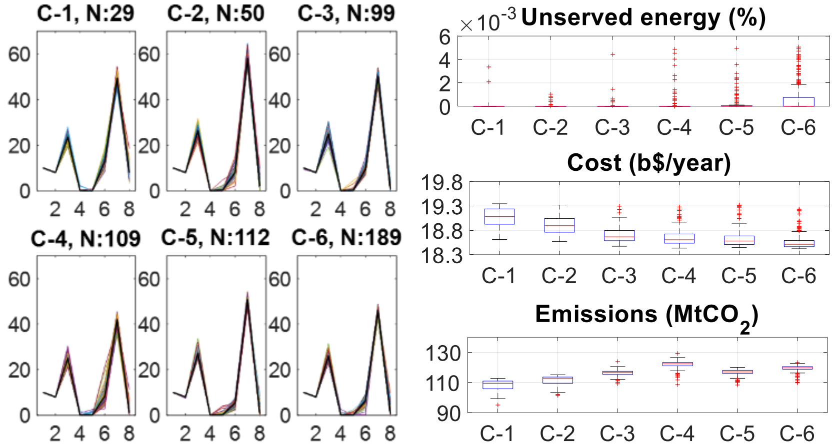


Fig. 10. (left) Six capacity clusters (y-axis = GW of capacity for each available generation option along the x-axis) and (right) unserved energy, cost and CO_2 emissions of all clusters CP35.

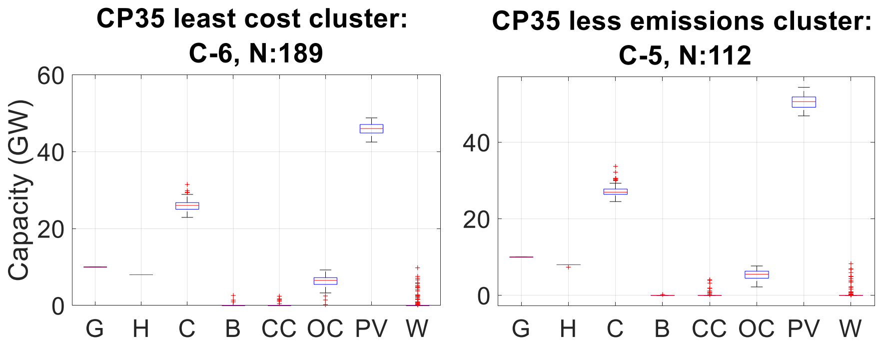


Fig. 11. (left) Technology mixes in the least cost cluster (C-6) and (right) technology mixes in the 'lower' emissions cluster (C-5) for 5% least cost portfolios with CP35.

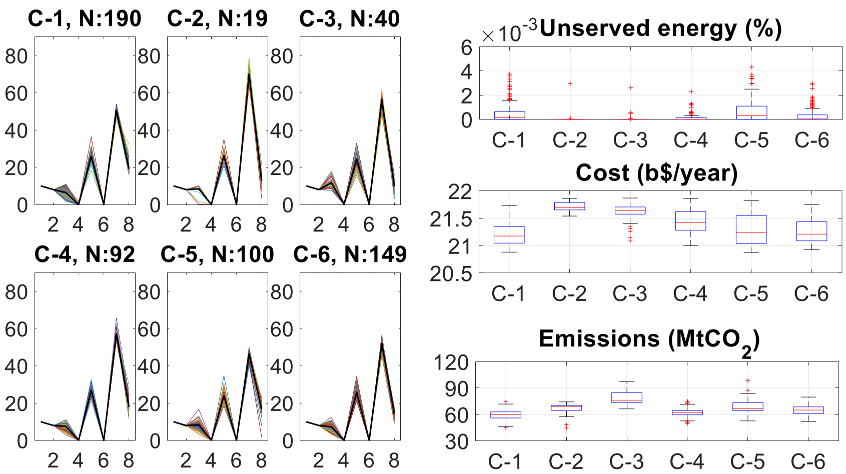


Fig. 12. (left) Six capacity clusters (y-axis = GW of capacity for each available generation option along the x-axis) and (right) unserved energy, cost and CO_2 emissions of all clusters CP60.

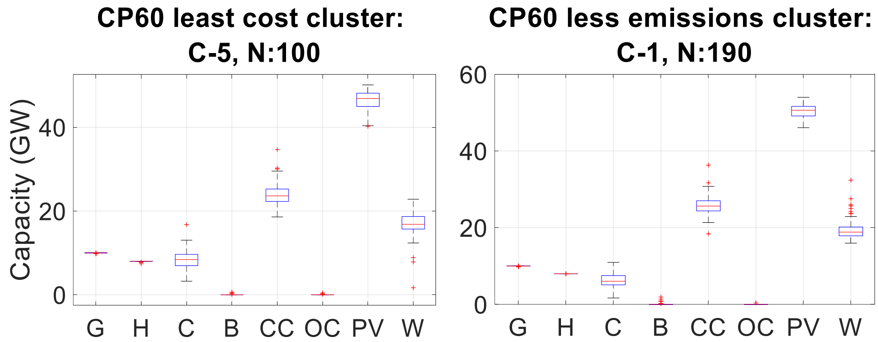


Fig. 13. (left) Technology mixes in the least cost cluster (C-5) and (right) technology mixes in the 'lower' emissions cluster (C-1) for 5% least cost portfolios with CP60 and fixed 0.005% USE limit.

While keeping the coal capacity range remain unchanged or with only little, insignificant variations, as shown in Fig. 9 (right), 11 (right) and 13 (right), higher PV capacities are deployed in all technology mixes in the less emissions cluster for all CPs compared to those shown in the least cost clusters. This increased PV deployment and hence reduced emissions comes at what would seem to be fairly low additional costs. The spread of the emissions and costs in the clusters also offers potentially valuable insights for policy makers in terms of what strategies driving particular technology deployment patterns might involve in terms of cost or emission risks.

3.5. Shared area of possible technology mixes

In Fig. 14 we present a method to compare possible technology mixes for each of the CP scenarios in Fig. 4. This visualization allows us to see the possible range of capacity deployment of each generation technology, and gain insights into the possible capacity trade-offs across both conventional and variable renewables.

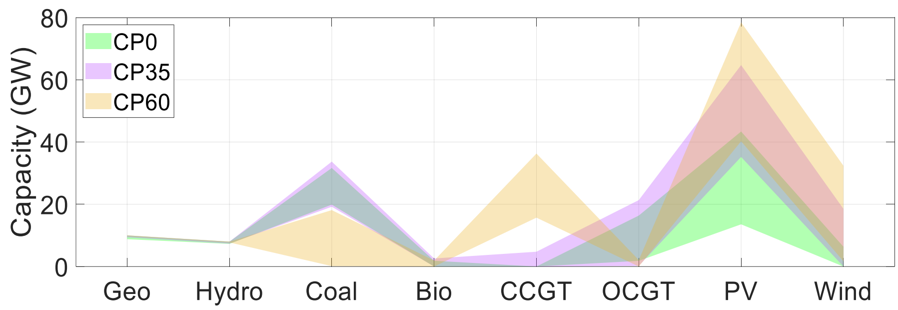


Fig. 13. Overlapping areas of the possible technology mixes of all CPs and its intersections.

It is interesting to note that the range of coal generation capacities in for CP0 and CP35 overlap almost entirely, despite more solar PV and wind capacity in CP35. A higher total generation cost in CP35 than in CP0 results from more capacity built, although coal fuel costs are reduced. CP60, by comparison, has low coal capacity and, a very different pattern of CCGT deployment, which sees almost no OCGT deployed. This reflects of course the higher emissions intensity but lower capital costs of OCGT versus CCGT. A higher carbon price increases the competitiveness of CCGT, even for low capacity factor operation.

The results indicate that if a relatively high social cost of carbon is considered, policy should be directed to deploy large-scale variable renewables penetrations, particularly solar PV. The results also indicate that limiting Java-Bali future coal capacity, in this case to less than 40 GW, it a low cost option either with or without consideration of the cost of carbon.

4. Concluding remarks and future work

Electricity industry planning for a more sustainable energy future is enormously challenging given the potentially conflicting objectives of affordability, security and environmental impacts, and the very high levels of future uncertainty regarding their prioritization, as well as in generation technology progress and costs. Developing countries face particular challenges in all these regards. While there is a wide and growing range of simulation and optimization tools to assist in such planning, they have inevitable limitations in terms of incorporating these uncertainties and mapping tradeoffs.

Our study sought to advance existing tools and methods in three key ways – identifying explicit metrics of the energy trilemma that could be used to assess different possible future electricity industry generation portfolios across all key dimensions, mapping the solution space of 'high performing' if not absolutely 'least cost' generation portfolios including through the application of clustering, and exploring ways to present the solution space to planners and policy makers.

We applied this to the question of possible future generation investment pathways for the Java-Bali interconnected system. Our techniques highlight the very diverse generation portfolios that delivered costs close to the 'least cost' portfolio. Given the uncertainties involved in these cost estimations, notably future technology and fuel costs, it is clear that policy makers and planners have a wide range of possible pathways towards a more sustainable electricity industry future. There are important tradeoffs, particularly between costs and emissions, in these choices. These techniques, as well as these particular findings, would seem to have wider relevance to the electricity industry planning and policymaking communities, especially in assisting them unveiling a wide range of possible generation portfolios mixes that are suitable to their preference pathways and electricity industries' context, given the techno-economic potential of high variable renewable energy penetrations.

On the other hand, the potential deployment of large-scale variable renewable energy technologies, has been unfortunately hindered by a range of barriers, including lack of supportive regulations and incentives for investment, and poor coordination between national and local government. The coal sector has to date received more support compared to large-scale renewables. The barriers of integrating solar and wind in Indonesia have been comprehensively presented and discussed in [46].

Given the range of potential generation portfolio options that could improve system reliability and offer a better environmental outcome at low cost, our results highlight the importance for policy makers in developing countries of considering a range of options to satisfy the trilemma objectives. This is particularly important given the uncertainties around technology costs.

While imposing a carbon tax on the supply side would be expected to increase electricity costs, our results show that a faster shift to high penetration renewables could in fact be relatively low cost, and therefore improve access and affordability with lower risk around the future costs of emissions and fossil fuels.

As indirect emissions such as those produced during solar cell manufacturing are insignificant compared to operational emissions from burning fossil fuels [47], we do not include these in our study. However, a more detailed study of environmental impacts for generation expansion planning could incorporate all life cycle pollutants. As the techniques presented here could be used to applied to these problems and a broader range of electricity industry planning problems where multiple objectives are to be satisfied, future work could also seek to address some of the limitations of the existing modelling framework including adding transmission costs and a greater role for different storage technologies, also to use more nuanced, weighted metrics of the trilemma.

Acknowledgement

This work is part of the study supported by funding from the Indonesia Endowment Fund for Education (LPDP) Scholarship provided by the government of the Republic of Indonesia under contract number: PRJ-3167/LPDP.3/2016. The authors thank PLN for providing 2015 Java-Bali hourly demand data.

References

[1] IEA, 2019. *World Energy Outlook 2019*, Paris: International Energy Agency.

[2] WEC, 2011. Policies for the future: 2011 assessment of country energy and climate policies. Accessed in 25 March 2019. Available online: https://www.worldenergy.org/publications/2011/policies-for-the-future-2011-assessment-of-country-energy-and-climate-policy/.

[3] WEC, 2018. World energy trilemma index 2018. Accessed in 30 March 2019. Available online: https://www.worldenergy.org/publications/2018/trilemma-report-2018/.

[4] Gunningham, N., 2013. Managing the energy trilemma: The case of Indonesia. *Energy Policy*, Volume. 54, pp. 184-193.

[5] IRENA, 2017. Renewable energy prospects: Indonesia, a REmap analysis. Accessed in 15 March 2018. Available online: https://www.irena.org/publications/2017/Mar/Renewable-Energy-Prospects-Indonesia.

[6] IRENA, 2017. Renewable power generation costs in 2017. Accessed in 14 March 2018. Available online: https://www.irena.org/publications/2018/Jan/Renewable-power-generation-costs-in-2017.

[7] Dagoumas, A. S. & Koltsaklis, N. E., 2019. Review of models for integrating renewable energy in the generation expansion planning. *Applied Energy*, Volume. 242, pp. 1573-1587.

[8] Ringkjob, H. K., Haugan, P. M. & Solbrekke, I. M., 2018. A review of modelling tools for energy and electricity systems with large shares of variable renewables. *Renewable and Sustainable Energy Reviews*, Volume. 96, pp. 440-459.

[9] Deng, X. & and Lv, T., 2020. Power system planning with increasing variable renewable energy: A review of optimization models. *Journal of Cleaner Production*, Volume 246, 118962.

[10] Zafeiratou, E. & and Spataru, C., 2018. Sustainable island power system – Scenario analysis for Crete under the energy trilemma index. *Sustainable Cities and Society*, Volume. 41, pp. 378-391.

[11] Handayani, K., Krozer, Y. & Filatova, T., 2017. Trade-offs between electrification and climate change mitigation: An analysis of the Java-Bali power system in Indonesia. *Applied Energy*, Volume 208, pp. 1020-1037.

[12] Vithayasrichareon, P. & MacGill, I. A., 2012.A Monte Carlo based decision-support tool for assessing generation portfolios in future carbon constraint electricity industries. *Energy Policy*, Volume. 41, pp. 374-392.

[13] DeLuque, I. & Shittu, E., 2019. Generation capacity expansion under demand, capacity factor and environmental policy uncertainties. *Computer & Industrial Engineering*, Volume 127, pp. 601-613.

[14] Santos, M. J., Ferreira, P. & Araújo, M., 2016. A methodology to incorporate risk and uncertainty in electricity power planning. *Energy*, Volume 115, pp. 1400-1411.

[15] PLN, 2019. *PLN Statistics 2018*, Jakarta: P.T. Perusahaan Listrik Negara (Persero).
[16] IESR, 2019. *A roadmap for Indonesia's power sector: How renewable energy can power Java-Bali and Sumatra*, Jakarta: Institute for Essential Services Reform.
[17] EMD International A/S, 2017. Wind energy resources in Indonesia. Accessed in 20 March 2018. Available online: http://indonesia.windprospecting.com/
[18] DEN, 2017. Technology data for Indonesian power sector: catalogue for generation and storage of electricity. Accessed in 2 March 2018. Available online: https://www.den.go.id/index.php/publikasi/download/55
[19] Elliston, B., MacGill, I. & Diesendorf, M., 2014. Comparing least cost scenarios for 100% renewable electricity with low emission fuel scenarios in the Australian National Electricity Market. *Renewable Energy*, Volume. 66, pp. 196-204.
[20] CEEM, 2018. National Electricity market Optimizer (NEMO). Accessed in 20 February 2018. Available online: http://ceem.unsw.edu.au/open-source-tools
[21] Tanoto, Y., MacGill, I., Bruce, A. & Haghdadi, N., 2017. *Photovoltaic deployment experience and technical potential in Indonesia's Java-Madura-Bali electricity grid*. Asia Pacific Solar Research Conference., 2017, pp. 1-13.
[22] Auger, A. & Hansen, N., 2012. *Tutorial CMA-ES: evolution strategies and covariance matrix adaptation.* 14th Annual Conference Companion on Genetic and Evolutionary Computation., 2012, pp. 827-848.
[23] DEN, 2016. Indonesia energy outlook 2016. Accessed in 3 August 2017. Available online: https://www.den.go.id/index.php/publikasi/download/34
[24] Pfenninger, S. & Staffell, I., 2016. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy*, Volume. 114, pp. 1251-1265.
[25] MEMR, 2019. P.T. PLN (Persero) Electricity provisioning planning 2019-2028 (in Bahasa Indonesia). Accessed in 30 March 2019. Available online: http://gatrik.esdm.go.id/assets/uploads/download_index/files/5b16d-kepmen-esdm-no.-39-k-20-mem-2019-tentang-pengesahan-ruptl-pt-pln-2019-2028.pdf
[26] Motlagh, O., Berry, A. & O'Neil, L., 2019. Clustering of residential electricity customers using load time series. *Applied Energy*, Volume. 237, pp. 11-24.
[27] McLoughlin, F. Duffy, A. & Conlon, M., 2015. A clustering approach to domestic electricity load profile characterization using smart metering data. *Applied Energy*, Volume. 141, pp. 190-199.
[28] Rhodes, J. D., Cole, W. J., Upshaw, C. R., Edgar, T. F. & Webber, M. E., 2014. Clustering analysis of residential electricity demand profiles. *Applied Energy*, Volume. 135, pp. 461-471.
[29] Satre-Meloy, A., Diakonova, M. & Grünewald, P., 2020. Cluster analysis and prediction of residential peak demand profiles using occupant activity data. *Applied Energy*, Volume. 260, 114246.
[30] Meschede, H., Esparcia Jr., E. A., Holzapfel, P., Bertheau, P., Ang, R. C., Blanco, A. C. & Ocon, J. D., 2019. On the transferability of smart energy systems on off-grid islands using cluster analysis – A case study for the Philippine archipelago. *Applied Energy*, Volume. 251, 113290.
[31] Zhou, Y., Lork, C., Li, W. T., Yuen, C. & Keow, Y. M., 2019. Benchmarking air-conditioning energy performance of residential rooms based on regression and clustering techniques. *Applied Energy*, Volume. 253, 113548.
[32] Scott, I. J., Carvalho, P. M. S., Botterud, A. & Silva, C. A., 2019. Clustering representative days for power systems generation expansion planning: Capturing the effects of variable renewables and energy storage. *Applied Energy*, Volume. 253, 113603.
[33] Pineda, S. & Morales, J. M., 2018. Chronological time-period clustering for optimal capacity expansion planning with storage. *IEEE Transaction on Power System*, Volume. 33 (6), pp. 7162-7170.

[34] Malik, A., Haghdadi, N., MacGill, I. & Ravishankar, J., 2019. Appliance level data analysis of summer demand reduction potential from residential air conditioner control. *Applied Energy*, Volume. 235, pp. 776-785.

[35] Kotu, V. & Deshpande, B., 2015. *Predictive Analytics and Data Mining: concepts and practice with Rapidminer, in Chapter 7 – clustering*. Waltham: Morgan Kaufmann.

[36] Mehrotra, K., Mohan, C. K. & Ranka, S., 1997. Elements of artificial neural networks. Cambridge: MIT Press.

[37] Llanos, J., Morales, R., Núñez, A., Hernández, R. & Lanas, F., 2017. Load estimation for microgrid planning based on a self-organizing map methodology. *Applied Soft Computing*, Volume. 53, pp. 323-335.

[38] Oprea, S. V. & Bâra, A., 2016. *Electricity load profile calculation using self-organizing maps.* 20th International Conference on System Theory, Control and Computing, 2016, pp. 860-865.

[39] Calinski, T. & Harabasz, J., 1974. A dendrite method for cluster analysis. *Comm. in Statistics*, Volume. 3, no. 1, pp. 1–27.

[40] Davies, D. L. & Bouldin, D. W., 1979. Cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume. 1 (2), pp. 224-227.

[41] Saitta, S. Raphael, B. & Smith, I. F. C., 2008. A comprehensive validity index for clustering. *Intelligent Data Analysis*, Volume. 12 (6), pp. 529-548.

[42] Liu, Y., Li, Z., Xiong, H., Gao, X. & Wu, J., 2010. *Understanding of internal clustering validation measures*. IEEE International Conference on Data Mining Workshops, 2010, pp. 911-916.

[43] Veldhuis, A. J. & Reinders, A. H. M. E., 2013. Reviewing the potential and cost-effectiveness of grid-connected solar PV in Indonesia on a provincial level. *Renewable and Sustainable Energy Reviews,* Volume. 27, pp. 315-324.

[44] Clark, R., Zucker, N. & Urpelainen, J., 2020. The future of coal-fired power generation in Southeast Asia. *Renewable and Sustainable Energy Reviews*, Volume. 121, 109650.

[45] MJHR, 2014. Government Regulation Number 79/2014 on the National Energy Policy. Accessed in 20 April 2020. Available online: https://policy.asiapacificenergy.org/node/3016.

[46] Burke, P. J., Widnyana, J., Anjum, Z., Aisbett, E. & Resosudarmo, B., 2019. Overcoming barriers to solar and wind energy adoption in two Asian giants: India and Indonesia. *Energy Policy*, Volume. 132, pp. 1216-1228.

[47] Kreith, F., Norton, P. & Brown, D., 1990. A comparison of CO_2 emissions from fossil and solar power plants in the United States. *Energy,* Volume. 15 (12), pp. 1181-1198.