New Methodology for Long-Term Transmission Grid Planning – General Description

Camille Pache, Jean Maeght, Baptiste Seguinot, Alessandro Zani, Sara Lumbreras, Andrés Ramos, Sergeï Agapoff, Leif Warland, Luis Rouco and Patrick Panciatici

***Abstract*—This paper presents a new methodology for long-term transmission planning over large systems. The developed approach aims at finding the optimal design of a large grid including its modular development plan over a long time horizon. Advanced optimization and simulation methods have been investigated to tackle this very large and complex problem, which includes highly combinatorial aspects and stochastic behaviours of system components, while ensuring some control over the system. Some tools have been implemented and tested on a case study based on the French and Spanish systems. The preliminary results of this study give an estimation of the resources that would be needed for a real study over a system of the size of the whole pan-European network for the period 2020 to 2050.**

I. Introduction

Long-term transmission planning over large power systems, such as the European transmission network, is a very large and complex problem. Indeed, transmission planning is usually performed for time horizons comprised between 10 and 20 years and current methods have not coped so far with planning over large areas and with increased uncertainty when integrating significant shares of renewable generation. Several studies are conducted on these topics [1]-[2]. Today, any national transmission system in Europe is planned using expert judgement. Therefore, Work Package 8 of the e-Highway2050 project [3] aims at defining a new methodology where long-term planning is formalised as an optimisation problem, and at specifying new tools. Advanced optimisation and simulation methods are investigated to tackle this complex highly combinatorial problem, taking into account the stochastic and dynamic behaviours of the system. A Monte Carlo approach has been retained to consider the stochasticity of the problem. The problem is solved from a transmission operator perspective, with no control over generation planning. Though, several possible evolutions of the electricity sector are investigated for different *scenarios*. The main challenges arise from the spatial, temporal and stochastic complexities. Thus, methods to reduce the size of the grid or to choose relevant snapshots among the 8760 hours of a year have been developed. The process followed by the developed methodology is represented in Fig. 1.

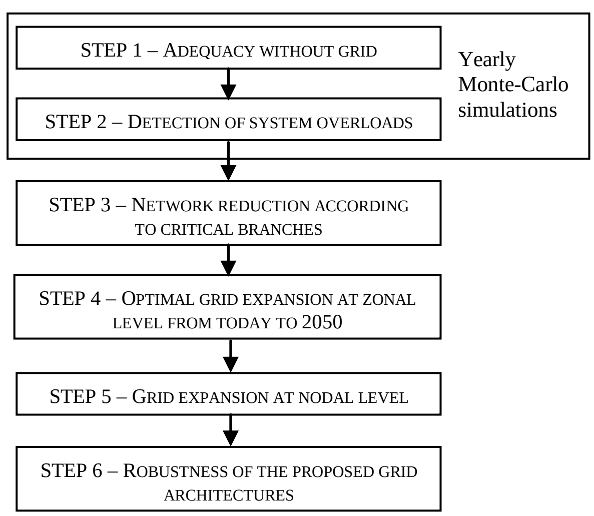

Fig. 1. Successive steps of the methodology

In a first step, controllable generation and consumption are calculated with an hourly time step to ensure power adequacy between production and load for each scenario (i.e. generation mix) and time horizon. Grid constraints are not taken into account at this step and several patterns of uncertainties are considered. In a second step, overload problems are detected on a simplified initial pan-European nodal grid (around 1,000 nodes). Then, in a third step, the initial network is reduced according to critical branches, leading to a zonal initial grid (100 nodes). In a fourth step, the modular development plan is calculated at the zonal level considering all time horizons and the whole set of scenarios. Starting from the zonal modular development plans, a grid expansion is performed for the first two time horizons at a nodal level. Finally, the robustness of the nodal grid architectures is checked to ensure that these grids can be operated without major voltage or stability issues.

II. Adequacy without grid

The first step of the methodology simulates the operation of the electrical system in "copperplate" (i.e. with infinite transmission capacities) and generates balanced time series of consumption and generation for different patterns of uncertainty. These time series are then used in the following processes as reference injections, which serve as a base to design the future network expansions. A Monte Carlo approach has been retained to take into account all the stochastic variables of the problem and to study the response of the power system to diverse possible events.

A. *Time series generation*

A first module samples time series of all uncertain inputs (wind and solar power generation, load, hydro inflows and outages of thermal units) based on intrinsic characteristics observed in past realizations, notably seasonality, autocorrelation, probability density function and cross-correlation, and assumptions about future trends. Time series are generated at an hourly time step, over a period of one year and for each area of the modelled power system.

A method has been proposed to learn the trend, seasonality and stochastic properties of the uncertain inputs from historical data. The residual and stochastic part of the process is described either by a seasonal ARMA [4] or a diffusion-type model [5]. Historical data have to be analyzed in order to select the most suitable model.

The inter-dependencies between time series are also handled by the time series generator. The impact of spatial correlations between time series has been investigated on a French test case at the horizon 2030 [6]. 100 Monte-Carlo years have been simulated with and without spatial correlations. The expectation of reliability indicators is reported in TABLE I. The main result of this test case is that the reliability of the system can be overestimated when spatial correlations are neglected.

A prototype has been implemented in Matlab.

TABLE I - Sensibility of reliability indicators to spatial correlations

	Energy not served (GWh/year)	***Loss of load duration (h/year)***	***Energy in excess (GWh/year)***	***Total cost (G€/year)***
With spatial correlations	194	42	33	8.86
Without spatial correlations	5	4	12	6.60

B. *Adequacy simulation*

The following step consists in simulating the behaviour of controllable generation and demand. First, two heuristics schedule the maintenance of thermal units and allocate hydro resources among the 52 weeks of the year. This weekly decomposition is solved with a rolling-planning which considers the approximate forecast of the uncertain inputs.

Finally, detailed adequacy simulations with an hourly time step are performed for each week to compute the hourly hydro-thermal dispatch. The retained adequacy model includes a detailed representation of the thermal units, with unit commitment (UC), minimum stable power and minimum on-line and off-line durations. The unit commitment model is based on the efficient model proposed in [7]. Moreover, consumption can be controlled via Demand Response (DR) programs, which allow to shift a part of the load within a given delay time. Hydro units and exchanges with neighbouring systems are modelled as well.

For each scenario, each time horizon and each Monte Carlo year, the adequacy simulator returns the time series of generation and consumption, as well as the availability of each thermal unit and the system cost. A prototype has been implemented in AMPL [8] with FicoXpress solver [9].

III. Detection of system overloads

The initial pan-European grid is first reduced to get a simplified nodal grid of around 1000 nodes. Then, overload problems are detected starting from the generation and consumption data previously generated.

A. *Network reduction*

In this subtask, a new clustering approach for a large transmission network was developed in order to reduce the initial existing European EHV grid (around 8000 nodes) to 1000 nodes. The resulting "nodal grid" will be the most detailed grid modelling used in the following steps. Two existing methods have been studied: k-means algorithm [10] and Dodu's mixed integer linear programming [11]. Each of them has some disadvantages, such as the resulting interfaces between two zones might not be physically correct; or the fact that the number of zones is difficult to control. Hence, we propose a new hybrid method which combines both methods and ensures that the resulting delineations are consistent with physical connections and allow some control over the number of zones [12]. Electrical distance is used to aggregate the nodes that are electrically close to each other. Additional rules were also introduced in order to take into account Phase-Shifting Transformers (PST), High-Voltage Direct Current (HVDC) links embedded in AC grid and cross border lines. A prototype has been implemented in Python.

B. *Automatic mapping*

Time series of load and production generated by the adequacy simulations for each country are disaggregated on each node of the simplified nodal grid. This is performed according to pre-defined distribution keys for all types of load and generation, and for each scenario and time horizon.

C. Detection of overload problems

Detecting overload problems on a large grid becomes challenging since time complexity is added to spatial complexity. Indeed, generation and consumption are defined with hourly time steps for a large number of Monte Carlo years, for each scenario and time horizon, leading to millions of injection patterns. One of the challenges is to detect which of these patterns will lead to congestions. We want to solve this problem without selecting ex-ante some representative snapshots, which does not ensure an exhaustive detection of overload problems, especially in a system with increasing uncertainties. In order to detect overload problems, the flows on the simplified nodal grid are computed for all injections using DC Optimal Power Flows (DCOPF) while considering the possible network controls (HVDC, PST).

Each DCOPF is modelled as a linear optimization problem in which controllable injections, voltage angles, PST phase angles and controllable flows over HVDC links are variables; it aims at minizing the cost of adjusting controllable generation and load computed in Step 1, while complying with network constraints and using in the best way costless controls (HVDC, PST). Then, different indicators on branches and nodes are derived from the DCOPF outputs and used to identify critical branches, as explained in the next section.

This DCOPF model has been written in AMPL [8] and solved with FicoXpress [9].

IV. Network reduction according to critical branches

Critical branches are defined as the transmission lines that present a special interest for the purposes of transmission expansion planning, notably lines which are frequently congested. Different indicators have been identified to select those branches: average flow above a certain threshold percentage of the line capacity, proportion of operation hours where the line is congested, dual value of the maximum flow constraint of the line, and the difference between marginal nodal prices at the end nodes of the line. Given the level of flow control they provide, HVDC lines or PSTs are always considered as critical branches.

Once critical branches are identified, a heuristic algorithm builds an initial network partition based on these critical lines. This initial partition is then refined using a modified version of the k-means. We propose to base the network partition on a composite distance which aggregates electrical distances and additional information relevant to transmission planning (e.g. geographical distances or marginal nodal prices). The network partition assigns nodes to zones in order to obtain a unique set of zones, while preserving critical branches whose end nodes must belong to different zones. This approach is illustrated in the case study presented in section VIII. Then, network reduction calculates the zonal network parameters (namely, the capacities and reactances of the inter-zonal corridors) that approximate the characteristics of the original system as accurately as possible.

V. Optimal expansion at zonal level from today to 2050

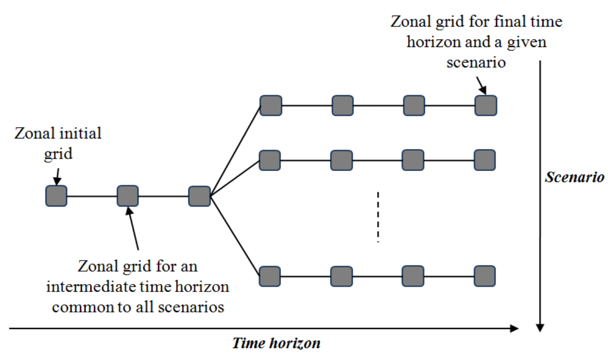

Fig. 2. Expansion planning over the time period 2020-2050

The goal of this step is to provide a sequential Transmission Expansion Plan up to 2050, with a common development of the network for all the scenarios up to 2030 (see Fig. 2). The expansion plan is given for the zonal network (100 nodes) and aims at minimizing both the investment cost of the grid and the operation and maintenance costs of the transmission system (the operational cost mirrors the costs related to congestions). Three different architectures will be considered and compared: an optimal solution (minimization of the sum of investment, operational and maintenance costs), a “super-grid” solution (penalization of short distance branches) and a “local development” solution (penalization of long distance lines).

A. Snapshot selection

Transmission Expansion Planning (TEP) is usually performed on a few operating situations or snapshots. However, two or three snapshots cannot be representative of a whole year especially in systems with a high penetration of renewable energy sources. Ideally, all the time steps and Monte Carlo years would be considered to optimize the grid development, which would require huge computing power. Thus, we suggest to select a subset of snapshots as representative as possible and use it to assess the reliability and the profitability of the transmission grid. The proposed approach is based on k-means and uses criteria drawn from Monte Carlo simulations to find representative snapshots.

Because of their variability, non controllable demand and generation are good indicators of variations in the system. Having in mind that the selected snapshots will be used in an investment planning for the transmission grid, marginal price differences between each node are considered as well.

B. Candidate selection

In order to reduce the computation time of the TEP optimization, the number of candidates (possible new transmission lines) proposed to the problem should be reduced. This module is based on the assessment of the profitability of each candidate [13] in a given state of the grid. A first selection is applied according to technical constraints, such as length or nature of crossed area. Then, potential candidates are iteratively identified and installed in the grid through simplified TEP optimizations (relaxing integer variables). It allows the algorithm to find candidates which become profitable once others have already been installed. Finally, a candidate analysis is performed to identify complementary and substitute candidates.

C. Transmission expansion planning optimization

The purpose of our method is to decide in each scenario and time-horizon which candidates are the most beneficial for the system. The cost involved by the investment made for each candidate is assessed, as well as the operational consequences following this reinforcement. The generation and consumption computed in Step 1 are considered the ideal situation which minimizes system cost. Thus, one part of the optimization problem aims at minimizing the deviation from these ideal injections. Once an investment has been decided, we model the operational impact as the difference between the costs of operating the system with and without grid constraints. Therefore, an investment is profitable if its investment cost and its impact on operational cost are lower than the operational consequences of not investing.

The developed approach is based on two optimization levels: investment and operational levels. The objective of the investment problem is to minimize the overall cost all along the time horizons (2020 to 2050) and for all the scenarios (weighted by their probability of occurrence), where a part of the operating cost is itself an optimization (operational level) minimizing, for a given grid and a selected set of snapshots, the production and consumption deviations from Step 1. We formulated this problem as a large Mixed Integer Linear Program (MILP) using an extended disjunctive model [14], in the DC approximation. It is implemented in AMPL [8].

VI. Grid expansion at nodal level

Once the three different modular development plans have been established on the zonal grid (100 nodes), the resulting architectures are defined on the simplified nodal grid (1000 nodes) for the first two time horizons. They should comply with the grid capacities defined in the zonal approach and ensure (N-1) system reliability taking into account all possible flexible devices.

For each inter-zonal capacity, several candidates are selected between each pair of zones using heuristics. Then, an optimizer chooses the best configuration mixing all the candidates for all the inter-zonal capacities, ensuring reliability and fulfilling the given inter-zonal capacities.

VII. Robustness of the grid architectures proposed

In the final step, the robustness of the proposed nodal grid architectures for the first two time horizons is checked to ensure that these grids can be operated without major voltage or stability concerns. This analysis involves the study of voltage-reactive power control and stability, transient stability and small-signal stability. However, we need to develop an AC load flow model of the grid to account for reactive power flows and bus voltage magnitude variation. Moreover, transient and small-signal stability analyses not only need the steady-state model of the network provided by an AC load flow, but also the dynamic models of generators and other dynamic devices. Thus, simplified models of dynamic components (e.g. wind generator and VSC-HVDC link) are also developed in this step.

A method to build a fully detailed AC load flow from its DC approximation model data has been developed. However, the main issue of this method was the convergence of the initial load flow which could occur in a planning study and does not occur on actual realized data. If the initial AC load flow does not converge, a new algorithm using first order sensibilities of maximum mismatches has been proposed and implemented. It operates on shunt devices or reactive load control to guarantee AC load flow convergence and has been tested in two models of the French and the Spanish power systems. The performance of the algorithm is presented in

TABLE **II** and

TABLE **III**.

As models of wind generators and HVDC links are very complicated, their fundamental aspects have been identified to simplify them. The impact of their control schemes on the system performance was studied on a two-area system with wind generators and one HVDC link of LCC and VSC technologies. The conclusion is that control schemes of wind generators and HVDC links have a great impact on the critical clearing time, and thus on the system performance.

TABLE II - SHUNT DEVICE CONTROL REQUIREMENTS AS FUNCTION OF LOAD POWER FACTOR

Load PF	0.989	0.980	0.950	0.900	0.850	0.801
QL [MVAr]	17420	23651	38282	56410	72183	87050
B [MVAr]	0	0	0	300	2052	11641

TABLE III - REACTIVE LOAD CONTROL REQUIREMENTS AS FUNCTION OF LOAD POWER FACTOR

Load PF	0.989	0.980	0.950	0.900	0.850	0.801
QLinitial [MVAr]	17420	23651	38282	56410	72183	87050
QLfinal [MVAr]	17420	23651	38282	56277	71368	80448
DeltaQL [MVAr]	0	0	0	-133	-815	-6602

Voltage-reactive power analysis aims at allocating reactive power compensation and voltage control resource in such a way that the voltage stability of the system is ensured. Starting from the converged AC load flow, a sequential linear programming optimization algorithm based on the first order sensitivities determines the reactive power compensation resources (reactors and capacitors) and the voltage control means (generator terminal voltages and transformer taps) to reach a desired voltage profile. Moreover, the distance to the voltage collapse is determined using a continuation load flow algorithm.

VIII. CASE STUDY

In parallel of the development and implementation of each step, modules are successively integrated and tested. The integration test is based on the current French-Spanish transmission system, including 2196 nodes and 3715 transmission lines. As the size of this network is close to the 1000-node simplified nodal network from the methodology, it did not have to be reduced. Then, starting from the 2012 installed capacities in both countries, 2 different scenarios for 3 time horizons (2030 to 2050) were developed. The first scenario keeps the same energy mix than today, while the second has a more important part of renewable energy sources; 10 Monte-Carlo years are run per scenario and time horizon. The system includes around 300 thermal units.

The first modules which have been integrated were the time series generator, the adequacy simulator, the disaggregation of time series and the DCOPFs on the full network.

All the simulations are run on a dedicated Linux server composed of 16 cores. Parallelisation on the different cores is used to speed up the process. The computation times obtained for the adequacy simulations and DCOPFs over one year using FicoXpress [9] are presented in Table IV.

This gives us an estimation of the resources that would be needed for a larger study. If we were to simulate 2000 Monte Carlo years for 7 different scenarios and 6 time horizons, the run of these modules would take around 400 days of computation on 160 cores, creating 500TB of output data.

TABLE IV - COMPUTATION TIMES AND DATA VOLUMES FOR ONE SIMULATED YEAR (2040, SCENARIO 1, 6 YEARS IN PARALLEL)

Module	*Time*	*Data volumes*
Time Series Generator	1min	3.3 GB
Hydro Scheduling	10s	
Adequacy: for 52 weekly MILP problems	19min	
Disaggregation of Time Series	27min	
DCOPFs : for 8736 hourly linear problems	1h 22min	1.3 GB
Average running time per year	2h 9min	5.6 GB

Once all the hourly DCOPFs have been run, the average values and deciles of all indicators (DCOPF outputs) are calculated for each branch and each node over all scenarios and time horizons. These indicators are then used to select critical branches. Only 400kV nodes (542 nodes) are considered in the network partition process; the other nodes will be assigned to the closest 400kV node. In this preliminary case study, the top 100 lines by average flow were selected, in addition to another 17 special elements identified as critical as well. The initial partition resulted in 24 zones. Then, the partition was refined into 50 zones using a composite distance (50% electrical distance and 50% geographical distance); they are displayed in Fig. 3. The resulting partition using a composite distance is more compact geographically than if only the electrical distance is used.

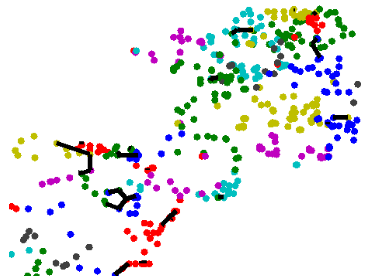

Fig. 3. Network reduction for the case study using a composite distance

The next step will be the integration of the next modules and their application to this case study. The full methodology and associated results will be available at the end of the e-Highway2050 project, in December 2015.

IX. References

[1] Y. Li, and J.D. McCalley, "Design of A High Capacity Inter-Regional Transmission Overlay for the U.S.," *IEEE Trans. Power Systems*, vol. PP, pp. 1-9, Jun. 2014.
[2] F.D. Munoz, B.F. Hobbs, and S. Kasina, "Efficient proactive transmission planning to accommodate renewables," in *2012 IEEE Power and Energy Society General Meeting*, pp. 1-7.
[3] *e-Highway2050*, http://www.e-highway2050.eu
[4] D.C. Hill, D. McMillan, K.R.W. Bell, and D. Inflied, "Application of Auto-regressive Models to UK Wind Speed Data for Power System Impact Studies," *IEEE Trans. Sustainable Energy*, vol.3, pp. 134-141, Jan. 2012.
[5] M. Doquet, "Use of a stochastic process to sample wind power curves in planning studies," in *Power Tech 2007 IEEE Lausanne Conf.*, pp.663-670.
[6] *Generation Adequacy Report 2012*, RTE, Paris, France. [Online]. Available: http://www.rte-france.com/uploads/Mediatheque_docs/vie_systeme/annuelles/bilan_previsionnel/an/generation_adequacy_report_2012.pdf
[7] G. Morales-España, J.M. Latorre, and A. Ramos, "Tight and Compact MILP Formulation of the Thermal Unit Commitment Problem," *IEEE Trans. Power Systems*, vol. 28, pp. 4897-4908, Nov.2013.
[8] *AMPL Optimization*, http://ampl.com.
[9] FICO Xpress Optimization Suite, *FICO*, http://www.fico.com/en/ products/fico-xpress-optimization-suite/.
[10] N. Shi, Liu X., and Y. Guan, "Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm, *" 2010 Third International Symposium Conf. on Intelligent Information Technology and Security Informatics (IITSI)*, pp. 63-67.
[11] J.C. Dodu, J. P. Ludot, and J. Pouget, "Sur le regroupement optimal des sommets dans un réseau électrique," *RAIRO, Recherche opérationnelle*, vol. 3, pp. 17-37, 1969.
[12] C.-A. Fezeu, K. Bell, J. Ding, P. Panciatici, and M.-S. Debry, "Simplified representation of a large transmission network for use in long-term expansion planning," presented at the PSCC Conf., Wroclaw, Poland, 2014.
[13] S. Lumbreras, A. Ramos, and P. Sánchez, "Automatic selection of candidate investments for Transmission Expansion Planning," *International Journal of Electrical Power & Energy Systems*, vol. 59, pp. 130–140, Jul. 2014
[14] Bahiense, L., et al., "A mixed integer disjunctive model for transmission network expansion," *IEEE Trans. Power Systems*, vol. 16, pp. 560-565, Aug. 2001.

X. Biographies

Camille Pache graduated in Electrical Engineering at *École Supérieure d'Electricité (Supélec)*, Paris, France (2013). She is working in the R&D department of RTE (French Transmission System Operator), Versailles, France.

Jean Maeght received a M.Sc. in Mathematics (1997) from Orsay and Toulouse Universities, France, and a Ph.D. in mathematical optimization (Toulouse University, 2000). He is working in the R&D department of RTE.

Baptiste Seguinot received an Engineering degree from *Ecole Centrale de Lyon*, Lyon, France (2012). He is working in the R&D department of RTE.

Alessandro Zani graduated in Informatics Engineering (2006), received a M.Sc. in Electricity Sector from Universidad Pontificia Comillas, Madrid (2008) and a Ph.D. in computational methods (2011). He is working in RSE.

Sara Lumbreras received a M.Sc. in Industrial Engineering and obtained a Ph.D. in Transmission Expansion Planning (2014) from Universidad Pontifica Comillas, Madrid, Spain. She is working at Instituto de Investigación Tecnológica (IIT), Madrid, and is lecturer at the Industrial Organization Department at ICAI.

Andres Ramos graduated in Electrical Engineering (1982) and obtained a Ph.D. in Electrical Engineering (1990) from Universidad Politécnica de Madrid, Spain. He is a Research Fellow at IIT, and Professor at Comillas' School of Engineering (Head of the Department of Industrial Organization).

Sergeï Agapoff graduated in Electrical Engineering at *École Supérieure d'Electricité (Supélec)*, Paris, France (2012). He is working at RTE.

Leif Warland received a M.Sc. in Electrical Engineering at Norwegian University of Science and Technology (NTNU), Trondheim, Norway (1996) and a Ph.D. in Electrical Engineering (NTNU, 2001). He is working at SINTEF, Trondheim, Norway.

Luis Rouco graduated in Industrial Engineering (1985) and obtained a Ph.D. in Industrial Engineering (1990) from Universidad Politécnica de Madrid. He is a Research Fellow at IIT and Professor in the Department of Electrical Engineering at Comillas' School of Engineering. Member of CIGRE, IEEE.

Patrick Panciatici graduated in Electrical Engineering at *École Supérieure d'Electricité (Supélec)*, Paris, France (1984). He is working at RTE, as scientific advisor in the R&D department. Member of CIGRE, IEEE, SEE.