

Systematical Determination of Load Flow Cases for Power System Planning

G. Rechberger, H. Renner, A. Gaun

Abstract-- This paper presents a novel method to determine a minimum number of representative load flow cases with load and generation data in a meshed transmission grid, which cover all critical power line loading situations in a given time frame. Measured active and reactive line loading data obtained from SCADA are the basic data input for this method. In order to find the optimum number of load flow cases a heuristic search algorithm in combination with a genetic optimization algorithm is used. The comparison of the required load flow cases with conventional load situations e.g. peak load or peak generation show, that conventional load flow cases do not cover all critical power line loading situations. Load flow cases determined with the presented method give an objective and representative picture of the expected situations and can be used for power system planning in the view of increasing demand and new generation capacity installations.

Index Terms-- load flow case optimization – power flow measurement – power system planning – genetic algorithm – set of load flow cases

I. INTRODUCTION

THE measured time series of reactive and active power, usually available in 15 min resolution, concerning consumption, generation and loading of lines provide a huge amount of data representing the state of an electrical power system. Based on this data set, a single load flow case for each time interval can be extracted. These load flow cases, consisting of load and generation values, represent a base for load flow analyses considering future demand increase, installation of new power plants and transmission network expansions for future situations.

In [2] peak load and peak generation cases for exemplary summer and winter snapshots are used as the planning bases. Other methods apply typical snap shot data sets for the planning purposes [6]. These situations are not necessarily representative for the maximal loading of each power line, especially in case of systems with high wind penetration, pump storage power plants or extensive cross boarder energy trading. In [7] a random set of load flow cases from a one year set of measurements with hourly resolution is recommended. The number of load flow cases in the dataset makes this

method highly computationally inefficient, requiring considerable effort concerning handling and utilization. The usage of selected load flow cases, with respect to maximal loading of a specific power line, provides a solution for this problem.

In this paper, a novel methodology for finding a minimum set of load flow cases based on active and reactive power flow measurements by the SCADA system is presented. The most straightforward method determines a set of load flow cases containing specific load flow cases for every worst-case bi-directional power flow scenario of a single power line. This set is minimized by a combination of a heuristic search algorithm and an optimization algorithm and covers all relevant critical loading situations of power lines.

II. METHOD

A. Data Bases

The introduced method is based on the measured apparent power S_{meas} with the elements $S_{meas,t,i}$ with the columns i , depending on the number of lines L in the investigated grid area and T rows for each single time step (index t) of the interested period T .

With respect to a worst case approach to the bi-directional power flows on the power lines the information about the load flow direction in the matrix S is necessary. For example, a planned power plant can cause additional line power flows in various directions, which have to be added or subtracted from the actual values of a load flow case. Since the apparent power S_{meas} has no information about the load flow direction, a direction (sign) of the active power is assigned to the matrix S_{meas} which results in the matrix S as follows:

$$S_{t,i} = S_{meas,t,i} \cdot sign(P_{t,i}) \quad (1)$$

where $P_{t,i}$ are the measured active power flows on the power lines.

With the given maximal allowed currents by the vector I_{lim} and the elements $I_{lim,i}$ of each power line, the conventional thermal transmission capacities $S_{lim,i}$ and the loadings $s_{t,i}$ in p.u. can be calculated (2).

Here is assumed that the actual voltage does not considerably deviate from the rated voltages U_i in the vector U where it is also possible to use the actual voltage at the maximal measured apparent power.

In the following line loadings are expressed in p.u. according to (2) and named with lower case characters.

G. Rechberger is with the Institute of Electrical Power Systems, Graz, University of Technology, A-8010 Graz (e-mail: georg.rechberger@tugraz.at)

H. Renner is with the Institute of Electrical Power Systems, Graz University of Technology, A-8010 Graz (e-mail: herwig.renner@tugraz.at)

A. Gaun is with Institute of Electrical Power Systems, Graz University of Technology, A-8010 Graz (e-mail: alexander.gaun@tugraz.at)

$$S_{lim,i} = \sqrt{3} \cdot U_i \cdot I_{lim,i} \quad \text{and} \quad s_{t,i} = \frac{S_{t,i}}{S_{lim,i}} \quad (2)$$

The separate measurement for each circuit of double circuit lines with the same start and end node for each circuit can be reduced to a single value.

The data corresponding to any irregular states of the grid e.g. failure, planned outage or special grid configurations due to congestion management, have to be considered separately. These irregular states require separate corresponding grid states for further load flow calculations.

B. Defined minimum percentage of analyzed load flows

When analyzing the network, it is useful to take into account only power lines with a maximal loading above a specified loading limit s_{base} . With $s_{base} = 20\%$ of their installed capacity all lines in the grid with peak loading values exceeding 20% are taken into account.

With $s_{base} = 60\%$ only lines with peak loading over 60% of the maximal transmission capacity of this line are evaluated. If an increase of loading, caused by load increase or new power plants, a $s_{base} = 60\%$ is not useful because in meshed grids loading over 60% are often not within the (n-1) security criteria [6]. Choosing s_{base} should regard the expected loading increase within the actual planning period. To estimate the loading increase of single power lines load flow calculations or the calculation of power distribution transfer factors PDTF [8] have been proposed.

The matrix $S_{t,i}$ (1) can be thus reduced by eliminating the 'non critical' power lines with maximal loading below the predefined percentage s_{base} to the vectors s_{max} and s_{min} with the elements $s_{max,i}$ and $s_{min,i}$ by the following equations:

$$\begin{aligned} s_{max,i} &= s_{t,i} \forall \max_t(s_{t,i}) \geq s_{base} \\ s_{min,i} &= s_{t,i} \forall \min_t(s_{t,i}) \leq -s_{base} \end{aligned} \quad (3)$$

Every line with loading situation exceeding s_{base} or below $-s_{base}$ is classified in further analyses as critical situation.

C. Critical range

Depending on the maximum in both load flow directions of the apparent power ($s_{max,i}$ and $s_{min,i}$) on every power line a critical range ϵ expressed as percentage of the loading limit in each direction can be defined (see (4)). Each load flow case with a line loading within the critical range ϵ is assumed to a representative peak loading case for this line. The critical range is described with the lower and the upper apparent power $s_{lower,i}$ and $s_{upper,i}$.

With $\epsilon = 1\%$ nearly all peak loading situations within the time period of measurement are obtained, with $\epsilon = 20\%$ deviations up to 20% from the selected load flow cases to the peak load situations are expected. Another possibility is to derive the critical range from the duration curve for the loading of each line. For further proceeding the definition for the critical range in (4) is used.

In Fig. 1 an example of load and duration line of the load

flow on a high voltage power line illustrates the critical range. In this example the critical range ϵ is defined with $\epsilon = 20\%$.

$$\begin{aligned} s_{upper,i} &= s_{max,i} - s_{lim,i} \cdot \frac{\epsilon}{100} \\ s_{lower,i} &= s_{min,i} + s_{lim,i} \cdot \frac{\epsilon}{100} \end{aligned} \quad (4)$$

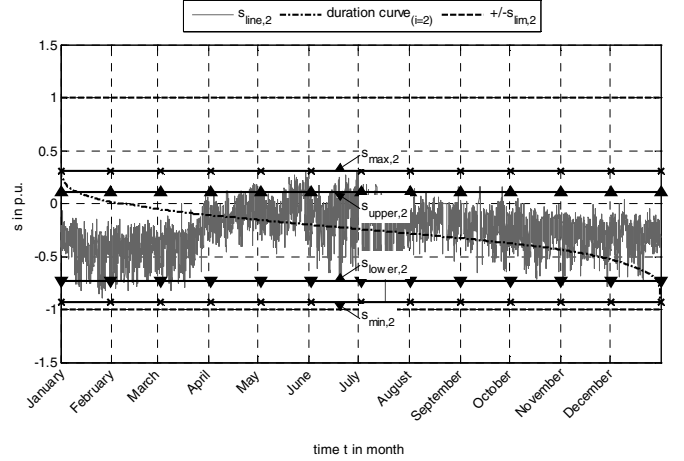


Fig. 1 Power transmission line ($i=2$) apparent power $s_{line,2}$ in p.u., duration curve, lower and upper critical area with $\epsilon = 20\%$.

A general survey of the general load flow situation for 10 typical lines chosen from the studied dataset in the interesting period is given in Fig. 2. This kind of diagram of the load flow situation used also in the following shows only the first 10 of the 38 considered power lines for the sake of a clear arrangement. This figure shows the measured line loadings $s_{line,t,i}$ including the maximal and minimal values $s_{line,i,max}$ and $s_{line,i,min}$. The critical range is between $s_{upper,i}$ and $s_{line,i,max}$ for one load flow direction and between $s_{lower,i}$ and $s_{line,i,min}$ for the opposite load flow direction. The dotted lines mark the s_{base} percentage.

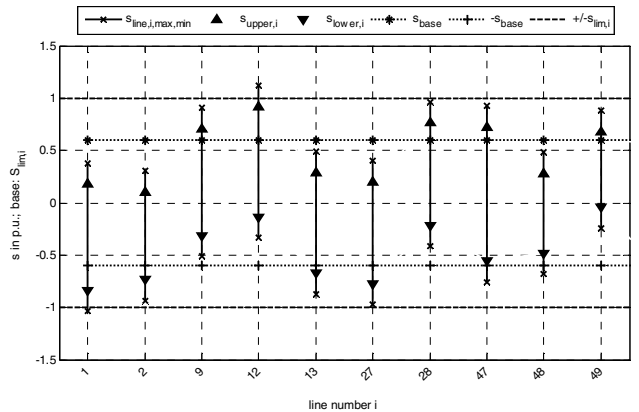


Fig. 2 Load flow situation of 10 power lines in a meshed grid (critical range for $\epsilon = 20\%$, $s_{base} = 60\%$).

All load flow values $S_{max,t,i}$ and $S_{min,t,i}$ of the remaining lines are tested whether they represent a peak loading situation in accordance to the above definition of the critical range for one or more lines. The results of the mentioned comparison are the logic suitability matrices C_{max} and C_{min} with entries for

every time step and every line.

$$C_{\min,t,i} = \begin{cases} 1, & \forall s_{\min,t,i} < s_{i,\text{lower}} \\ 0, & \forall s_{\min,t,i} > s_{i,\text{lower}} \end{cases} \quad (5)$$

$$C_{\max,t,i} = \begin{cases} 1, & \forall s_{\max,t,i} > s_{i,\text{upper}} \\ 0, & \forall s_{\max,t,i} < s_{i,\text{upper}} \end{cases} \quad (6)$$

The matrix C_{\max} for one load flow direction merged with C_{\min} for the other load flow direction merged to the matrix C_0 as input data for further optimization.

$$C_0 = [C_{\max} \quad C_{\min}] \quad (7)$$

Each row in matrix C_0 represents a single load flow case and each column a load flow direction of a single line with peak loadings exceeding s_{base} . A logical 1 as matrix element declares the according load flow case as peak loading case for this line. The optimization problem is to find a minimum set of load flow cases, covering all peak loading situations in each direction. This selection can be done as described below.

D. Minimizing the data set

To reduce computational time, the matrix C_0 is minimized before further steps to the matrix C . Firstly all rows of the matrix C_0 , with the same characteristic, identified by all none zeros entries at the same position, can be reduced to a single row. The row with the maximal mean loading of this set with same characteristic will be selected for further calculations. Secondly all rows with only zero entries representing non critical load flow cases also have to be eliminated to reduce unnecessary calculations. This remaining set of load flow cases represented by the matrix C has at least one critical loaded line for each remaining time interval.

E. Selection of load flow cases

The trivial solution is to use two separated load flow cases for each line selected with the criteria e.g. maximal loading in both directions. In this case the number of required load flow cases N_{LC} is equal to $N_{LC} = 2 \cdot L$. This builds a complete worst case set of load flow cases. The following method allows the reduction of required load flow cases to minimize computation time on the one hand and improve the handling of complex planning calculations on the other hand.

In step 1 the sum of the non-zero entries of each column (sc) of C is calculated. If C includes a column with only a single non-zero entry, this row respectively this load flow case is selected because this peak loading situation for this line is not contained in any other situation. This row represents the first selected load flow case with the time $t_{LC,k=1} = t_1$ (see (8)) and the first generated solution vector $C_{\text{comb}} = C_{t_1}$. In step 2 the matrix C will be reduced by excluding the row C_{t_k} and all columns with none zero entries of C_{t_k} .

Step 1 and step 2 will be repeated as long as there are single none zeros entries in the columns of the new reduced matrix C .

$$t_k = t \left[\text{first} \left(\sum_{i=1 \dots T, i} (C) = 1 \right) \right] \quad (8)$$

If no more single critical situation is found, in step 3 the sum of the non-zero entries of each row (sr) of C is calculated. The next selected load flow case with the time t_k (see (9)) is the row with the maximal sum over the rows. Now the matrix C will be reduced again as pointed out in step 2.

$$t_k = t \left[\max \left(\sum_{i=1 \dots N} (C) \right) \right] \quad (9)$$

Thus it is possible to find a new time t_k and the corresponding load flow case C_{t_k} with step 1 or step 3. The new solution vector C_{comb} is the result of the combination of the old solution vector C_{comb} and C_{t_k} . The OR disjunction of the two rows C_{comb} and C_{t_k} result in a new solution vector C_{comb} .

Step 1, 2 and 3 will be repeated as long as the solution vector C_{comb} is fully occupied. Now the time vector t_{LC} has an entry for each required load flow case.

$$t_{LC} = [t_1 \dots t_k] \quad (10)$$

The described method to determine a set of load flow cases which are required to represent all critical situations in a grid and a given period is recapitulated in the pseudo code diagram shown in Fig. 3.

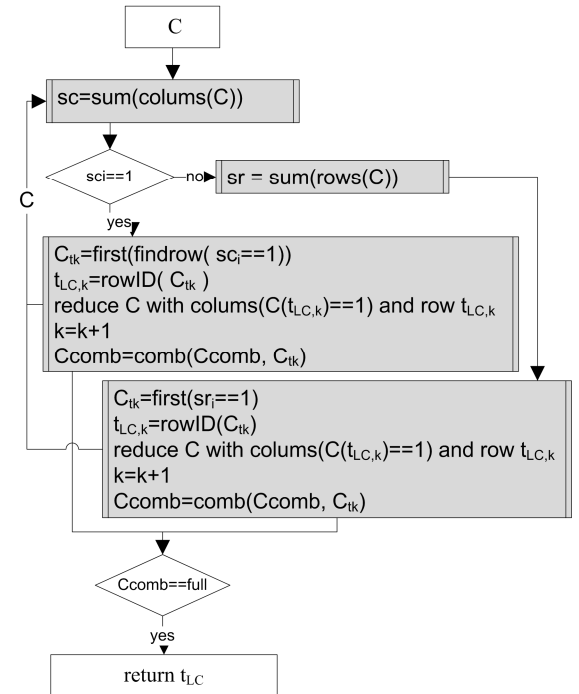


Fig. 3 Selection of load flow cases, search algorithm – pseudo code flow chart.

With this heuristic approach of selecting load flow cases, it is possible to find a minimum number of load flow cases required to express a maximum loading situation within the critical area for each line. Although this number of load flow cases is a local minimum but it is not guaranteed to be a global minimum. For small grids and few measured load flow cases,

it is possible to check every combination of the rows of C and find the optimal solution, but with measurement of e.g. one year and $\frac{1}{4}$ -hour resolution values it is not possible to solve this problem efficiently. A further problem is to select a load flow case in step 1 if there are more columns with single critical situations or in step 3 if the summation results in more than one maximum with the same value. In both cases this method selects the first detected entry.

F. Genetic algorithm

To improve the heuristic approach which mostly results in a local minimum number of load flow cases a genetic algorithm [1] is applied. In case of two or more load flow cases, represented by rows of the matrix C , with the same number of critical situations characterized by a non-zero entry at each respective positions we get a problem selecting a load flow case with the target to find a minimum number of required load flow cases. The heuristic approach, described above, will select always the first load flow case of the set of cases with the same maximal number of critical situations.

The implemented genetic algorithm includes the main variables Individual, Population, Parents, Children, Crossover rate, Permutation rate, Mutation rate, Number of iterations and Population size. These variables are used in the genetic algorithm consisting of the fitness function, the selection (two round tournament) and the reproduction (crossover, permutation and mutation).

An individual as part of a population is the modified order of the time T . The start population and the following populations consist of single individuals with random respectively by the genetic algorithm modified orders of the time T . The improved selection of load flow cases starts with sorting the rows of C in the order of the individuals of the population vector in step 1.

The number of individuals of populations describes the population size (popsize).

The fitness is the number of required load flow cases, resulting from the load flow cases selection. This is commonly known as fitness function.

The selection of individuals for the next population is done with a two rounds tournament selection which is based on minimum numbers of load flow cases as result of the fitness function. These individuals make up the parents of the new population. The two rounds of tournament selection require a population size which is divisible by four.

The crossover swaps parts of the parents individuals of the new population. The random determined crossover points define the parts of the parents. The crossover probability (crossover rate) was set to 0.5.

The permutation changes a random number of elements in individuals of the new population consisting of parents and children. The permutation probability (permutation rate) is also 0.5.

The mutation is the randomized modification of single elements in the population with mutation probability defined by a mutation rate of 0.5.

The individuals are characterized by the order of the time

steps. For crossover, permutation and mutation thus it is necessary to change the order in the individuals and not simply the elements and parts of the individuals.

Selection and reproduction (crossover, permutation and mutation) are repeated as long as a fixed number of iterations n_{iter} is reached.

G. Method overview

In Fig. 4 the implementation of the method to determine a set of load flow cases including the selection of load flow cases and the genetic algorithm is shown.

The result of the method, the time vector t_{LC} including the times of the selected load flow cases provides a set of load flow cases including load and generation values which are the bases for further load flow calculations in power system planning.

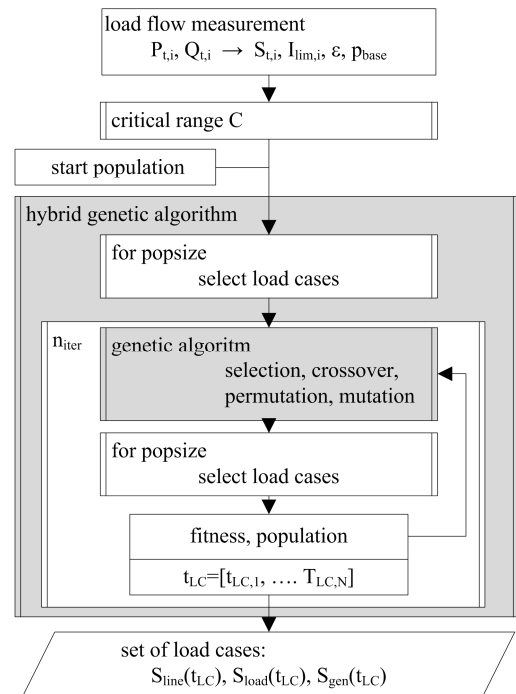


Fig. 4 Hybrid genetic algorithm¹ in combination with the selection of load flow cases – pseudo code flow chart.

III. RESULTS

A. Test Grid I and Grid II

The application of the introduced method is demonstrated with two real high voltage grids (GRID I and GRID II). For both grids the measured active and reactive power is based on one year observation with a resolution of $\frac{1}{4}$ hours and includes values for each power line, each generation unit and each load.

The high voltage GRID I consists of 107 power lines. This grid is linked to the surrounding network at three points. GRID II consists of 124 power lines and is linked to the surrounding network at four points. Both power systems include also generation (thermal, run of and pump storage hydro) and load (urban, industrial and rural).

¹ Hybrid genetic algorithm uses the local search algorithm to calculate the fitness during the evolutionary cycles from a population to the next one.

B. Population size and number of iterations

The coherence of required number of load flow cases depending on the number of iterations is investigated for different values of ϵ and s_{base} . After convergence of the fitness the genetic algorithm will be interrupted.

The population size was varied from minimum four to 200 (approximately the double number of lines) and shows also the same fast convergence. In most cases, when varying these parameters, the convergence value of the fitness function is already included in the initial population.

C. Set of load flow cases depending to the grid structure

To apply and evaluate the method of the determination of required load flow cases a set of different ϵ and s_{base} are used. The critical range was chosen as $\epsilon = 1\%$, 5% , 10% , 15% and 20% . In order to determine whether a single line should be evaluated or not $s_{base} = 20\%$, 40% , 50% and 60% are used.

Fig. 5 shows the determined, required 14 load flow cases for $\epsilon = 20\%$ and $s_{base} = 60\%$ of Grid I. For these parameters the analyses of the full data set S_{meas} with $L = 107$ power lines results in 38 power lines under consideration with 44 critical situations. In the diagram display $s_{LC,k}$ the loading situations of the considered power lines. The 14 determined load flow cases have at least one situation within the critical range between $s_{i,upper/lower}$ and $s_{line,i,max/min}$ of each of the 38 power lines. The enveloping lines s_e are the minimal and maximal values of the set of required load flow cases. The lines for $s_{LC,k}$ and s_e in the diagram represent only the time interrelationship of a load flow case and not a linear function between 2 power lines.

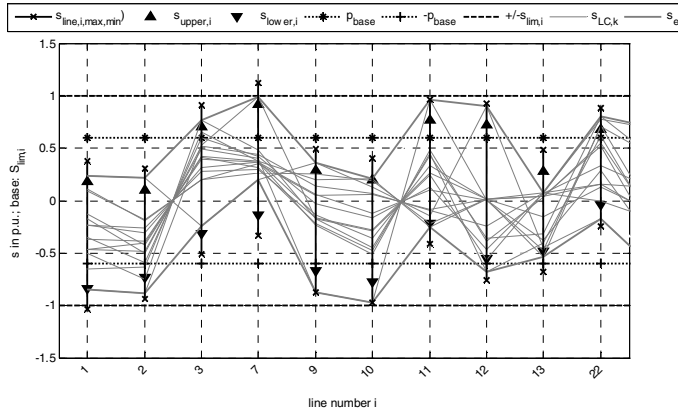


Fig. 5 Grid I, $\epsilon = 20\%$ and $s_{base} = 60\%$ - 14 required load flow cases.

Table I gives a summary of the required load flow cases for different choices of ϵ and s_{base} . The number of the considered power lines L_k as function of limit loading s_{base} is also listed in Table I. The number of critical situations CS is the number of lines with loading situations within the critical range in both directions. For different grids the number of required load cases N_{LC} is also different. N_{LC} for different ϵ and s_{base} of Grid II (Table II) have the same tendency as Grid I and the absolute values are higher compared to Grid I. The biggest effect on this increase is the higher number of lines in Grid II.

TABLE I
GRID I / GRID II - REQUIRED NUMBER OF LOAD FLOW CASES $N_{LC,I} / N_{LC,II}$ AND CONSIDERED NUMBER OF LINES (L_k) FOR ϵ AND s_{BASE}

s_{base}	20 %	40 %	50 %	60 %
$L_{k,I} / CS_I$	81 / 135	74 / 100	60 / 70	38 / 44
$L_{k,II} / CS_{II}$	115 / 170	88 / 104	71 / 83	50 / 54
$\epsilon = 1\%$	71 / 88	53 / 57	42 / 44	30 / 28
$\epsilon = 5\%$	47 / 56	38 / 37	34 / 33	26 / 22
$\epsilon = 10\%$	37 / 38	30 / 28	27 / 26	22 / 19
$\epsilon = 15\%$	30 / 32	26 / 25	24 / 21	18 / 17
$\epsilon = 20\%$	24 / 26	20 / 21	18 / 20	14 / 16

D. Grid separation

Grid II can be separated according to geographical boundary conditions in two parts in the following part A and part B. Part A includes 50 and part B includes 74 of the 124 power lines in Grid II. For both parts the required load flow cases are determined for the same parameters ϵ and s_{base} as claimed above. The results (number of load flow cases) are summarized in table III.

TABLE III
GRID II - PART A / PART B - REQUIRED NUMBER OF LOAD FLOW CASES $N_{LC,A} / N_{LC,B}$ AND CONSIDERED NUMBER OF LINES FOR ϵ AND s_{BASE} .

s_{base}	20 %	40 %	50 %	60 %
$L_{k,A} / CS_A$	48 / 75	38 / 47	30 / 36	22 / 24
$L_{k,B} / CS_B$	67 / 95	50 / 57	41 / 57	28 / 30
$\epsilon = 1\%$	47 / 45	30 / 30	21 / 25	14 / 15
$\epsilon = 5\%$	33 / 28	22 / 18	18 / 17	12 / 12
$\epsilon = 10\%$	24 / 21	18 / 15	15 / 14	12 / 9
$\epsilon = 15\%$	19 / 15	15 / 12	13 / 12	11 / 8
$\epsilon = 20\%$	16 / 13	13 / 11	13 / 11	10 / 8

The number of required load flow cases is also a function of the grid size and is related to the number of lines. The analysis of Grid II with $\epsilon = 1\%$ and $s_{base} = 20\%$ and with $L_k = 115$ lines and $CS = 170$ critical situations results in $N_{LC} = 88$ required load flow cases. Part A analyzed with same ϵ and s_{base} and with $L_{k,A} = 48$ and $CS_A = 75$ results in $N_{LC,A} = 47$. Part B with $L_{k,B} = 67$ and $CS_B = 95$ results in $N_{LC,B} = 45$. For smaller grids fewer required load flow cases are necessary but mostly the parts of a split grid requires for them alone more characteristic load flow cases as for the parts combined. The numbers of required load flow cases for the parts of a splitted grid also have a decreasing tendency as the complete grid depending to increasing p_{base} and ϵ .

Comparison peak load and peak generation case

On the hand of Grid I the power line loadings at peak load situation (Fig. 6) and at peak generation situation (Fig. 7) are compared with the critical range with $\epsilon = 20\%$ and $s_{base} = 60\%$.

The peak loading situation matches eight times the critical range. So it is possible to find 35 critical situations which are

not matching with the required criteria. As well the peak generation load flow case matches only eight times (different lines) the critical range.

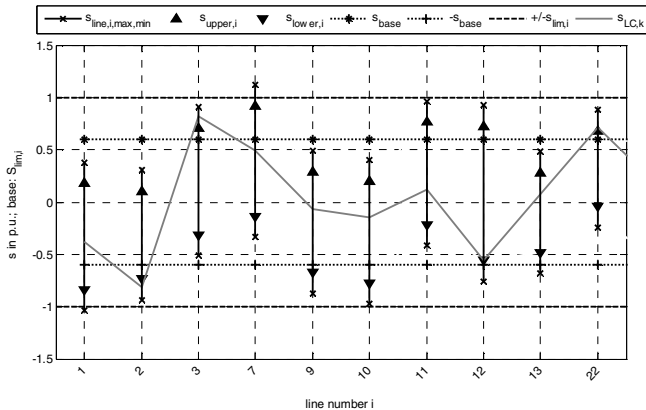


Fig. 6 Peak load situation $s_{L,C,k}$ and load flow situation of 10 power lines in a meshed Grid I; critical range with $\varepsilon = 20\%$ and $s_{base} = 60\%$.

The finding of these comparisons is that single load flow cases, e.g. peak load or peak generation, are not appropriate to give a worst case scenario of a power system. Most of the occurring critical situations are not taken into consideration with this single load flow cases.

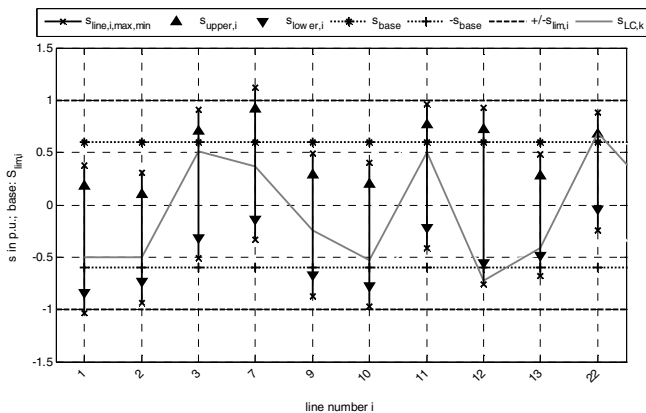


Fig. 7 Peak generation situation $s_{L,C,k}$ and load flow situation of 10 power lines in Grid I; critical range with $\varepsilon = 20\%$ and $s_{base} = 60\%$.

IV. CONCLUSION

With the presented method it is possible to find a minimum number of representative load flow cases which cover all critical power line loading situations in a given time frame.

The proposed method is primarily meant for use in meshed transmission grids and provides a systematical approach for power system planning.

For this method the bases data set, consisting of power flow measurement, is prepared with reducing double circuit lines, consideration of load flow direction, calculation line loadings, determination of a critical range of each power line, consideration only lines exceeding minimal loading and reducing of load flow cases with same characteristic.

The characteristic load flow cases are determined with a combination of a heuristic search algorithm and a genetic optimization algorithm to find a minimum number of load flow cases.

The number of required load flow case depends on the extension of the analyzed grid, on the claimed accuracy of the critical range ε and on the number of lines considered to be critical according to their maximal loading.

The classical planning scenarios e.g. peak load or peak generation usually only cover a limited number of critical loading situations. So they are not appropriate to give a worst case scenario of a power system.

V. REFERENCES

- [1] D. E. Goldberg, "Genetic algorithms for search, optimization, and machine learning," 1 edition, Boston, Addison-Wesley Professional, 1989
- [2] T.J. Hammons, "Configuring the power system and regional power system developments in Southeast Europe," in *Proc. 2004 Universities Power Engineering Conf.*, pp.1319 - 1326 vol. 2
- [3] G. Latorre, R. D. Cruz, J. M. Areiza, and A. Villegas "Classification of publications and models on transmission expansion planning," *IEEE Transaction on Power Systems*, vol. 18, no. 2, pp. 938-946, May 2003
- [4] D. Oeding, "Elektrische Kraftwerke und Netze," 6. edition, Berlin Heidelberg New York, Springer, 2004
- [5] *EURELECTIC WG SYSTINT, European Interconnection State of the Art 2002 (SYSTINT Annual Report)*, 2002-210-0002, 2002
- [6] *UCTE Operational Handbook – Appendix 4: Coordinated Operational Planning, capture B. UCTE network calculations*, v2.5 E, 24.06.2004
- [7] C. L. Krane, "Strukturbewertung elektrischer Übertragungsnetze," Dissertation, Dept. Electr. Power Syst. Univ. Aachen, 2007
- [8] Wood, Allen J., and Wollenberg, BruceF., "Power Generation, Operation and Control", 2nd edition, New York, Wiley-Interscience, 1984

VI. BIOGRAPHIES

Georg Rechberger was born in Linz, Austria, in 1978. He received his diploma degree in 2005 at Graz University of Technology, Graz, where he currently holds a position as scientific assistant at the Institute of Electrical Power System. His main work in research and teaching is in the field of electrical power system planning.

Herwig Renner was born in Graz, Austria, in 1965. He received his doctoral degree in 1995 at Graz University of Technology, Graz, where he currently holds a position as associate professor at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system analyses with special emphasis on quality, supply reliability and power system control and stability. In 2007 he was employed as visiting professor at Helsinki University of Technology.

Alexander Gaun was born in Kufstein, Austria, in 1979. He received his diploma degree in 2005 at Graz University of Technology, Graz, where he currently holds a position as scientific assistant at the Institute of Electrical Power Systems. His main work in research and teaching is in the field of electrical power system planning, reliability calculation, genetic algorithms and electromagnetic compatibility.