

# Principle component analysis of the spectroscopic and neutron parameters characterizing PF-1000 device plasmas

K. Jakubowska · W. Skrzeczanowski ·  
M. Paduch · M. Scholz · E. Zielińska

Received: 6 November 2013 / Accepted: 16 April 2014 / Published online: 10 May 2014  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** Principal Component Analysis has been employed in order to find correlations between different quantities (related to spectroscopy in visible range, neutron yield and the time derivative of the current) describing Z-pinch plasma in the PF-1000 installation. Set of measurements in various plasma conditions has been taken into account. Strong correlation between level of impurity content and neutron yield has been found.

## 1 Introduction

Principal component analysis (PCA) [1] has been used since many years in different areas of research (e.g., sociology, biology, economy and many other disciplines). The analysis leads to reduction in multidimensional arrangement of the data with implicit correlations with smaller number of variables. Usually, the results of this statistical procedure are presented in PCA plots representing measured variables in new coordinates (so-called principal components) that include as much as possible the variability of input data. This paper presents the first attempt of employing the PCA analysis of optical emission

in the visible range, neutron yield and differential current data in plasma-focus devices.

## 2 Experimental setup

The calculations are based on experiment, which has been carried out at the plasma-focus PF-1000 [2, 3] device in Warsaw, Poland. It consists of two coaxial electrodes of the same length (460 mm). The inner electrode (anode) is made of copper (diameter 230 mm). The outer electrode (cathode, diameter 400 mm) consists of 12 stainless steel tubes (diameter 82 mm). The anode is surrounded tightly by cylindrical, ceramic insulator (length 85 mm) positioned at the lower part of the anode. The gas breakdown develops along the external part of the insulator. The PF-1000 device is a Mather-type gas discharge facility in which plasma accelerates in the direction of the anode tube and after reaching the end of the tube turns around and implodes radially. The maximum plasma compression so-called *pinch* including almost total discharge current has the dimension of  $10 \times 1 \text{ cm}^2$  and is confined for around 100 ns. In most discharges, the capacitor bank, capable of accumulating 1,064 kJ energy at 40 kV, was charged to 24 kV, and in a few of them to 22 kV.

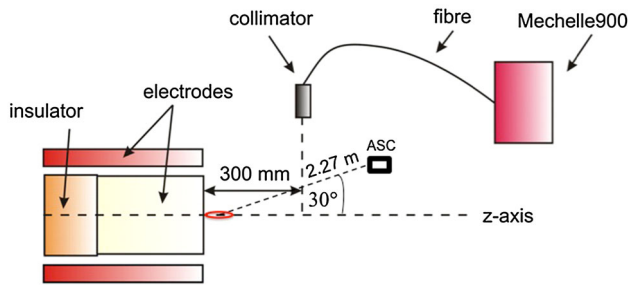
The experiment has been performed in vacuum chamber pumped out to the basic pressure of  $2 \times 10^{-5}$  hPa. The deuterium filling has been varied, for different discharges, in the range of 0.9–2.1 hPa.

The spectroscopic data have been acquired along the axis perpendicular to the z-axis. Spectrally resolved plasma radiation in the PF-1000 device was registered in the range of 350–1,000 nm with the use of Mechelle900 spectrometer and recorded with PCO SensiCam camera providing minimum time resolution of 100 ns.

K. Jakubowska (✉) · M. Paduch · E. Zielińska  
Institute of Plasma Physics and Laser Microfusion (IPPLM),  
01-497 Warsaw, Poland  
e-mail: katarzyna.jakubowska@ipplm.pl

W. Skrzeczanowski  
Military University of Technology, Institute of Optoelectronics,  
00-908 Warsaw, Poland

M. Scholz  
Institute of Nuclear Physics, Polish Academy of Sciences,  
Kraków, Poland



**Fig. 1** Geometry of the visible spectroscopy experimental setup including position of the silver activation counter position at the PF-1000

The neutron emission from PF-1000 discharges has been recorded by means of absolute calibrated silver activation counter. Geometry of the experiment is shown in Fig. 1.

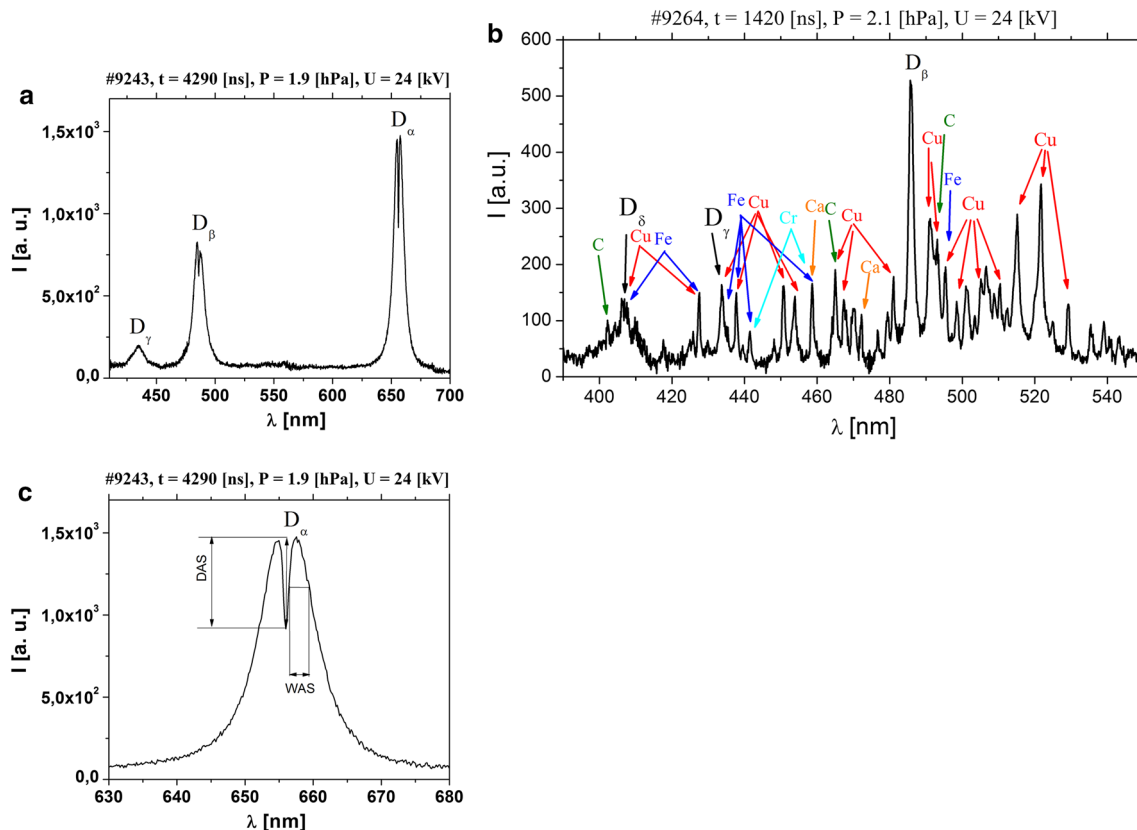
### 3 Experimental results

Typical PF-1000 spectrum measured in the visible region far from anode surface consists of Balmer series ( $D_\alpha$ ,  $D_\beta$  and  $D_\gamma$ ). In a very few cases, also  $D_\delta$  emission has been observed. In most of the discharges,  $D_\alpha$  and  $D_\beta$  emission has been highly self-absorbed, which is presented in

Fig. 2a. In some discharges, various impurity lines (mostly copper emission, e.g., CuI 510.554 nm, CuI 515.324 nm and CuI 521.820 nm) have been encountered. Detailed analysis of spectra implies presence of another impurity emission that can be attributed to the cathode and the insulator surface erosion (Fig. 2b). Unfortunately, resolution of Mechelle900 spectrometer is too low to distinguish and ultimately define impurity lines. As expected, Balmer series lines have been much broader than impurity lines, because of hydrogen, hydrogen isotopes and hydrogen-like ion emission are subjected to dominating Stark broadening. The neutron emission was of the order of  $1 \times 10^8 - 8 \times 10^{10}$  n/shot and the maximum neutron yield of  $1 \times 10^{11}$  measured with silver activation counter.

### 4 Discussion of the PCA analysis

Multivariate statistical analysis, the so-called PCA, was applied to find correlations between different parameters of plasma-focus discharge. It belongs to statistical factorial methods [1, 4–6] and is widely used in multivariable comparative analysis. This statistical approach allows comparing objects or sets consisting of many variables and



**Fig. 2** a Balmer series. b Example of the discharge with high amount of impurities, c the depth of the self-absorption peak (DAS) and the width of the right halves in self-absorption peak (WAS)

quickly finding the most important properties of these data. The method comes down to an orthogonal conversion of an input object set of any nature (input matrix) by a linear combination of new, non-correlated and sometimes non-observable characteristics called principal components. Multivariable comparative analysis leads to the following:

- Reduction in a high number of variables describing objects to a few fundamental ones which can be subject to detailed analysis;
- Grouping of objects characterized by similar properties;
- Separation of the most typical phenomena or processes, which leads to better preparation of experiments and reduces their costs;
- Clarification of relationships between object characteristics.

The incorrect selection of input variables can cause misleading results in comparative analysis. That is why the variables describing object set should be limited using substantial, formal and/or statistical criteria. The preliminary selection of variables based on substantial circumstances gives a distinction of potential statistical variables from which the number is reduced using statistical methods. This results in the formation of final set of variables, which are later used for comparative analysis. Our carefully selected experimental parameters were chosen as input variables.

The evident cost of reducing a high number of input variables, which represent the entire variability of input set (information included in the set), is the loss of a part of the information carried by lower number of principal components (two or three principal components). It is accepted that a representation of an investigated problem by principal components or factors is correct when it carries 70–80 % variability of the input set (sometimes, interpretation of a problem can be successful even for 50–60 % representation).

The number of principal components that should be retained in analysis has been defined by two criteria: Kaiser and Cattell (scree plot) criteria [1]. The Kaiser rule eliminates from further analysis all components, for which corresponding eigenvalues fall below unity. According to this criterion, only the first 3 components (explaining 69 % of variance) should be taken into account in the analysis of our data. The Cattell scree plot presents eigenvalues corresponding to different principal component numbers. Usually, the first component has the highest value, next decreasing gradually with the increase in the component number. It suggests neglecting all components that appear behind the one which starts the elbow in the scree plot.

The data analysis has been based on 74 discharges in various initial conditions (working gas pressure 1.0–2.93 hPa, capacitor bank charged to 22 or 24 kV). As

**Table 1** The parameters taken for analysis have been chosen arbitrary from much bigger matrix of various quantities (22 parameters)

Abbreviation	Parameter description
DC	The time of measurement in correlation with the moment when time derivative of the current (dI/dt) reaches minimum value
NY	The neutron yield
IC	The number of spectral lines in the wavelength range 490–530 nm (Fig. 2b)
IR	The intensity ratio of copper lines—CuI 521/CuI 515 nm
DAS	The depth of the self-absorption peak of $D_\alpha$ (Fig. 2c)
DBS	The depth of the self-absorption peak of $D_\beta$
WAS	The width of the right half in self-absorption peak $D_\alpha$ (Fig. 2c)
WBS	The width of the right half in self-absorption peak $D_\beta$
P	The pressure of working gas
CV	The capacitor bank voltage

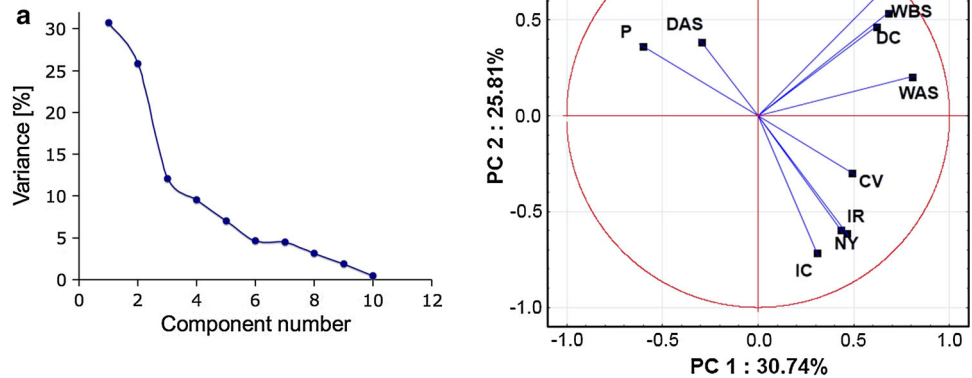
preliminary input variables, 10 parameters have been selected and taken into account (see Table 1, Fig. 2c).

The first three principal components (Fig. 3a) fulfill Kaiser criterion—their corresponding eigenvalues are  $>1$ , and they carry almost 70 % of input variability. At the same time, the level of variance explanation clearly reveals complexity of the physical phenomena and how difficult is to explain the PF discharge physics. Moreover, the relatively low level of variance amount carried by the first two principal components ( $<60$  %) for 10 input variables may be also due to a bit low population tested (as compared to variable number). Generally, the PCA results are more correct if the number of cases (observations) exceeds at least 7–10 (better 20) times the number of variables [1].

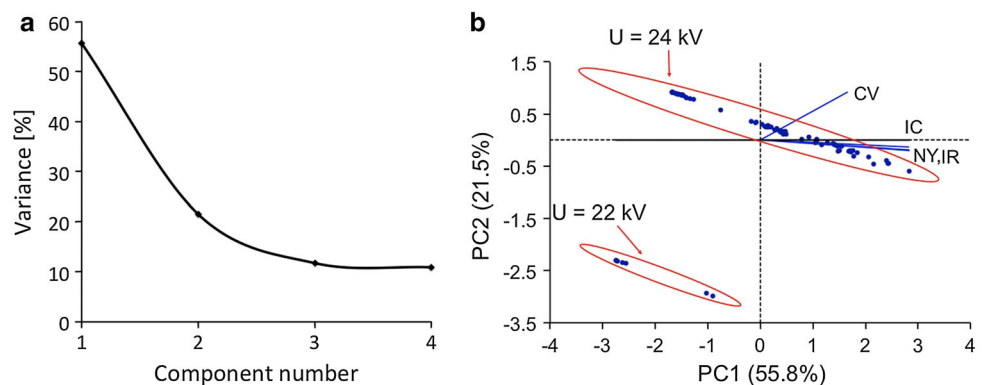
The results of the analysis are shown in Fig. 3b. The input variables are represented by vectors of lengths equal to the whole information attributed to the individual principal components (in our case to the first and second principal components). The closer an input variable is to the edge of the unit circle, the better its representation by principal components constituting coordinate system (i.e., the higher information amount carried by these PCs). If the whole information contributed by an input variable is carried by both PCs, the spot representing this variable (vector tip) would be positioned on the circle edge.

If the vectors (points) representing input variables (individual PF parameters) are close to each other, it means a strong positive correlation between those variables, while if the vectors are perpendicular to each other, any correlation between variables exists (the variables are non-correlated). For vectors directed oppositely and located almost on the same line, a very strong negative correlation between variables represented by these vectors is observed.

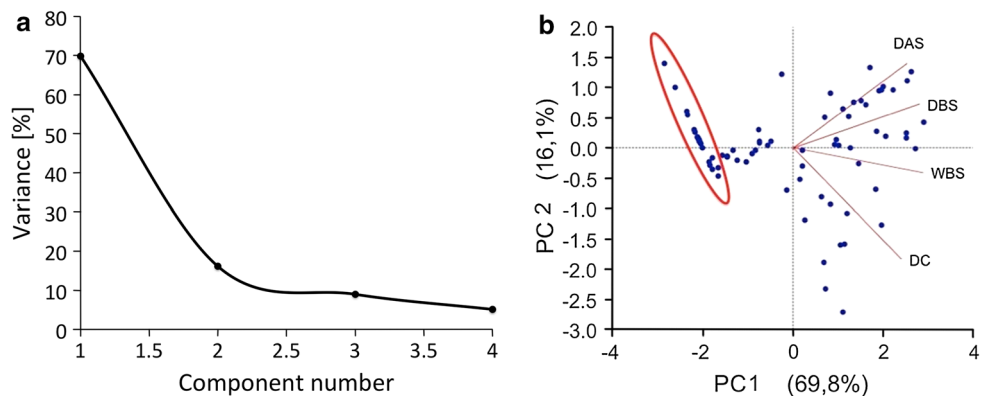
**Fig. 3** **a** The scree plot based on analysis of 10 parameters. **b** The principal component plot based on the first two PCs (PC1 and PC2)



**Fig. 4** **a** The scree plot based on analysis of the first group of parameters. **b** The PC plot based on the first two PCs (PC1 and PC2)



**Fig. 5** **a** The scree plot and **b** PC plot based on analysis of four parameters—the moment when  $dI/dt$  reaches minimum value (DC)—the depth of the self-absorption peak of  $D_{\beta}$  (DBS), the widths of the right halves in self-absorption peaks of  $D_{\alpha}$  (WAS) and  $D_{\beta}$  (WBS)



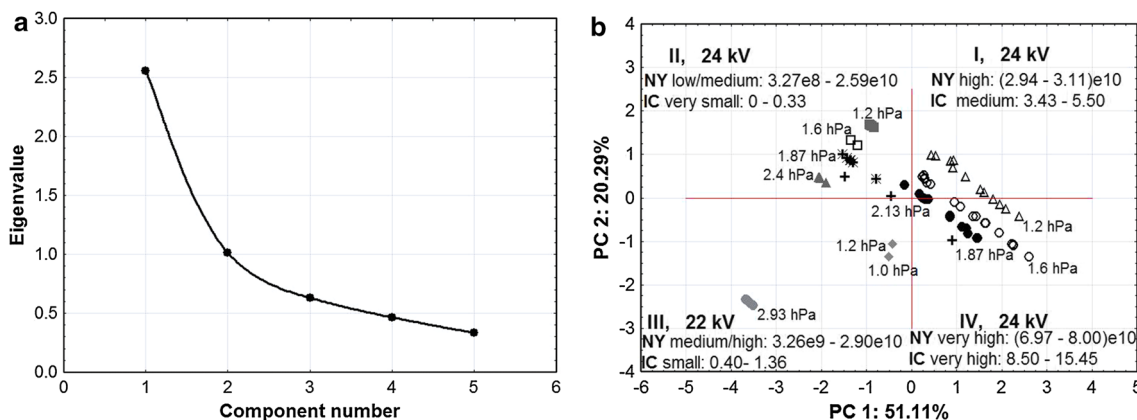
In our analysis, we used two software packages: StatistiXL (Figs. 3a, 4, 5, 6) [7] and Statistica 10 PL (Figs. 3b and 6) [8].

The principal component (PC) plot (Fig. 3b) showing correlations between 10 selected input variables in the first two principal components' (PC1 and PC2) plane very clearly indicates existence of three groups of quantities correlated with each other:

1. The first group consists of the neutron yield (NY), amount of impurity content in the wavelength range 490–530 nm (IC—defined as a number of lines

detected in wavelength range 490–530 nm), the ratio of copper lines amplitude—CuI 521/CuI 515 nm (IR) and the capacitor bank charge (CV);

2. The second group consists of the time of measurement in correlation with the moment when  $dI/dt$  reaches minimum value (DC), the depth of the self-absorption peak of  $D_{\beta}$  (DBS), the widths of the right halves in self-absorption peaks  $D_{\alpha}$  (WAS) and  $D_{\beta}$  (WBS) peak;
3. The third group consists of the depth of the self-absorption peak of  $D_{\alpha}$  (DAS) and the pressure of the working gas (P).



**Fig. 6** **a** The scree plot and **PC plot b** based on analysis of five parameters—neutron yield (NY), impurity content (IC), ratio of copper lines amplitude CuI 521/CuI 515 nm (IR), capacitor bank charge (P) and working gas pressure (CV)

The most evident correlation has been observed for neutron yield (NY) and the ratio of copper lines amplitude—CuI 521/CuI 515 nm (IR). It appears in the first group of the parameters. Performed PCA analysis has not confirmed correlation of the neutron level and deuterium pressure (Fig. 3b, [9]). The scree plot (Fig. 4a) analysis for variables belonging to this group of the parameters (four variables) implies that the first two principal components play the key role in explaining variance (77 %). Besides, the PCA of that group shows high correlation of the first principal component with all parameters confirming calculations executed for 10 parameters, explaining 56 % of variance and keeping positive sign for all of the parameters. The PC plot (Fig. 4b) drawn from the first and the second PCs confirms strong correlation between neutron yield and the intensity ratio of copper lines. It shows very clear linear relation between parameters separated in two groups’ discharges measured under capacitor bank charged to 22 and 24 kV.

The clear correlation between neutron yield (NY) and amount of impurities as well as the ratio of copper lines (CuI 521/CuI 515 nm) amplitude (IR) presented in Figs. 3b and 4b could be explained in three ways [3, 10, 11]:

1. Increase in the copper lines number can be related to the structure and intensity of the plasma current sheath formed on the anode surface. The high-density current sheath moving with high speed can release a big amount of copper from the anode before the plasma column is formed (pinch).
2. Radial collapse of the current sheath might be accompanied by magnetohydrodynamic (MHD) instabilities that can lead to disruption of the plasma column. As a result, high electric field can be observed at the position of disruption. The filamentation of the plasma leads to formation of electron and ion beams. The two

beams move along z-axis of the device. In this case, increase in the copper emission could be explained by the interaction of highly energetic beams with the anode material.

3. The strong correlation between neutron yield and amount of impurities most probably is related to emission of neutrons that are interacting with the surface of the electrode causing release and excitation of the electrode material.

The analysis performed for parameters belonging to the second group (DC, DAS, DBS and WBS, Fig. 5) does not indicate strong correlation between themselves, but it appoints a group of discharges that form straight line along which gathered points correspond to shots during which Balmer alpha emission has not been reabsorbed (inside the ellipse).

Similar PCA analysis was performed for five variables, i.e., neutron yield (NY), ion content (IC), intensity ratio of copper lines (IR), pressure (P) and capacitor bank charge (CV). Results, presented in Fig. 6, show projection of all investigated discharges on the PC plane. First two principal components carry together over 70 % of the input variability—their representation is satisfactory. Observed correlations between five input variables confirm relations found for four-variable case; high neutron yield and capacitor bank charge are accompanied by high ion content and ratio of copper emission (CuI 521/CuI 515 nm).

For better readability, in Fig. 6 are not marked values of CuI 521/CuI 515 nm line ratios (IR)—their behavior is very similar to that of impurity content (IC). Data presented in Fig. 6, separated with respect to the working gas pressure, are situated along straight lines, similar to the four-variable case shown in Fig. 4b where separation took place for capacitor bank charge.

## 5 Summary and conclusions

PCA has been employed in order to find correlations between optical spectra in the visible range, neutron yield and the time derivative of the current describing Z-pinch plasma in PF-1000 installation. Strong correlation between neutron yield and amount of impurities as well as the ratio of copper lines (CuI 521/CuI 515 nm) amplitude has been observed. It could be explained in three different ways:

- The thin current sheath moving with high speed can cause release of high amount of copper from the anode surface before the plasma column of high density is formed (pinch);
- The impurity atoms or ions are released from the electrode via interaction with energetic electron or ion beams;
- Most likely, impurity atoms or ions are created via interaction of the neutrons (being product of DD reaction) with the electrode material.

Performed PCA analysis has not confirmed correlation of the neutron level and deuterium pressure.

In order to draw reliable conclusions about the correlations and spectroscopy measurements itself (in particular, the evident strong correlation between neutron yield and impurity amount should be explained), it would be necessary to continue investigations with spectrometer providing higher spectral and time resolution with the help of a new arrangement of experimental chamber enabling the use of other optical diagnostics. This will be the next stage of our

work which will also be focused on searching further correlations between presented in the paper spectroscopic and other data and the parameters of X radiation measured in the PF-1000 device.

## References

1. I.T. Jolliffe, *Principal component analysis* (Springer, New York, 2002)
2. M. Scholz, R. Miklaszewski, V.A. Gribkov, F. Mezzetti, *Nukleonika* **45**(3), 155 (2000)
3. V.A. Gribkov, B. Bienkowska, M. Borowiecki, A.V. Dubrovsky, I. Ivanova-Stanik, L. Karpinski, R.A. Miklaszewski, M. Paduch, M. Scholz, K. Tomaszewski, *J. Phys. D Appl. Phys.* **40**, 1977 (2007)
4. R.A. Johnson, D.W. Wichern, *Applied multivariate statistical analysis* (Prentice Hall, Englewood Cliffs, 1992)
5. D.F. Morrison, *Multivariate statistical methods* (Mc Graw-Hill Book Company, New York, 1967)
6. N.H. Timm, *Applied multivariate analysis* (Springer, New York, 2002)
7. StatistiXL (<http://www.statiscixl.com>)
8. STATISTICA ver.10 (<http://www.statsoft.com/Products/STATISTICA-Features/Version-10>)
9. A. Kasperczuk, R. Kumar, R. Miklaszewski, M. Paduch, T. Pisarczyk, M. Scholz, K. Tomaszewski, *Phys. Scr.* **65**, 96 (2002)
10. P. Kubes, D. Klir, J. Kravarik, K. Rezac, J. Kortanek, V. Krauz, K. Mitrofanov, M. Paduch, M. Scholz, T. Pisarczyk, T. Chodukowski, Z. Kalinowska, L. Karpinski, E. Zielinska, *Plasma Phys. Control. Fusion* **55**, 035011 (2013)
11. M.J. Sadowski, M. Scholz, *Plasma Sources Sci. Technol.* **17**, 024001 (2008)