

# Vacuum leak search on Wendelstein W7-X using a numerical regression analysis

J. Baldzuhn<sup>\*</sup>, D. Rademann, H. Reimer<sup>1</sup>

Max-Planck Institute for Plasma Physics, 17491 Greifswald, Germany

<sup>1</sup> Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

<sup>\*</sup> Corresponding author. Tel: +49-3834-88-2534; fax: +49-3834-88-2509.

*E-mail address:* baldzuhn@ipp.mpg.de

## Abstract

The stellarator Wendelstein W7-X, a fusion plasma experiment, is almost ready to be set into plasma operation. Seventy superconducting coils produce the magnetic field that is required to confine the plasma. The coils are arranged in a cryostat vacuum vessel in shape of a torus. The superconductors are cooled by supercritical helium down to the nominal operating temperature between 3 K and 4 K. Before operation, vacuum leak search was performed on the machine, both the cryostat and the plasma vessel. Most of the leaks could be fixed; however not all. But the residual helium leak rate is too small to hamper the successful coil operation. A refined leak search is pending for the future, in order to localize and fix all the remaining leaks. This paper describes a numerical regression analysis, performed to localize these remaining leaks. The regression analysis could gain information on the size and position of the leak, and, at least in principle, on the hydraulic helium flow behavior inside the leak channel. It could be a helpful complement to the standard leak search techniques that had been used before.

**Keywords:** Vacuum leak search; Superconductors; Magnetic confinement; Regression analysis

## 1. Introduction and motivation

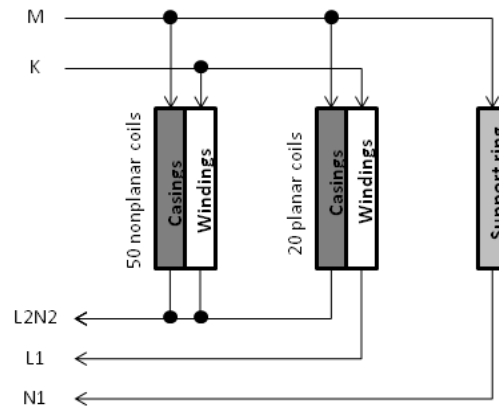
In this chapter we describe briefly the set-up of the Wendelstein W7-X cryostat, and the results of the standard leak search performed so far. A motivation is given for the refined leak search using the numerical regression analysis.

All superconducting coils for the stellarator W7-X [1,2,3] are situated inside a cryostat vacuum vessel [4]. Fifty of them are non-planar, twenty are planar coils. The set-up of the cryostat vessel, the arrangement of the plasma vessel, the vacuum systems and the standard leak search on them are described in detail in [5]. The technical commissioning of W7-X is described in [6]. In order to perform the standard vacuum leak search on the cryostat vessel, helium leak detectors were connected to the five vacuum pumping stations. That leak search was performed for leaks against air, the water cooling/heating pipes inside the cryostat, the helium supply pipes and all superconducting components. Helium gas puff was performed from the air side to localize leaks on the ports towards the experimental hall. The helium pressure inside the pipes was varied in time to record the time response of the leak detectors. Neon was applied in pressure steps to the water cooling pipes. Then mass spectrometers were used to measure leaking neon into the cryostat vacuum. The leak search was interrupted, after the remaining leak rate was considered as small enough to allow for the safe operation of the magnet system.

However, all pipes of the helium cooling system are connected to each other. No closing valves exist within the cryostat which can be used to separate the branches from each other. This is a consequence of the extremely narrow space situation inside the cryostat, and the requirement for a minimum number of feed-throughs for the valve actuators on the surface of the cryostat skin. In addition, each valve will increase the risk of leaks, the risk of a failure and it will increase the complexity of the system even more. For future machines it might be advisable to foresee, despite of these risks, some valves at strategic locations, for instance such that individual coils or groups of coils can be separated from each other. Grouping of the valves at positions close to the vessel wall is recommendable, if possible with a service flange close to them.

It is difficult to localize leaks within this distributed network. Nevertheless, helium pressure variations at the inlet pipes can show up to several seconds of delay time until they are measurable at more distant pipes, and sharp pressure transients at the inlet appear as temporarily smeared out response at the outlet. This hydraulic time response behavior of the entire pipe network is the key to the regression analysis. Fig. 1 shows a strongly simplified

sketch of the coil arrangement inside the cryostat, together with the five existing helium supply lines.

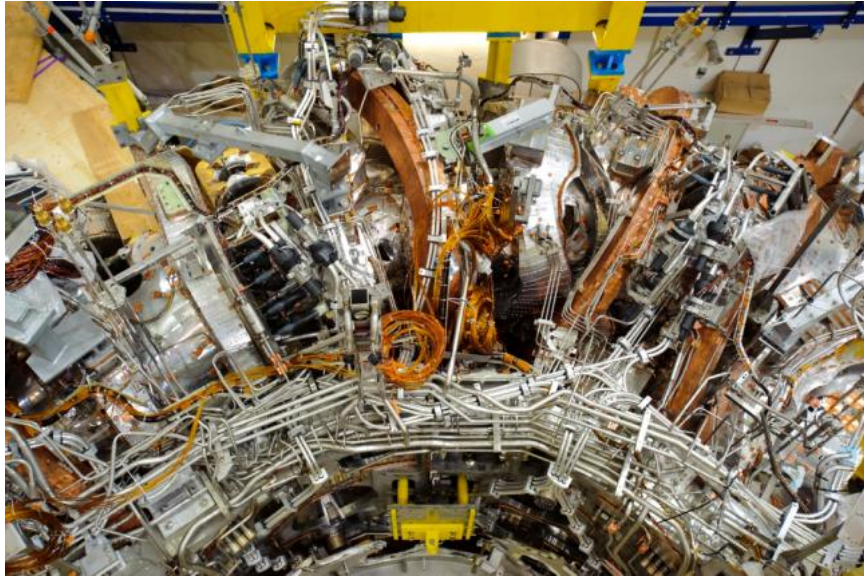


*Fig 1: Strongly simplified sketch of the superconducting coil arrangement inside the W7-X cryostat vessel. Shown are the helium supply/exhaust pipes labeled by “M”, “K”, “L2N2”, “L1”, “N1”, the sets of non-planar and planar coils including their windings and casings, and the coil support structure.*

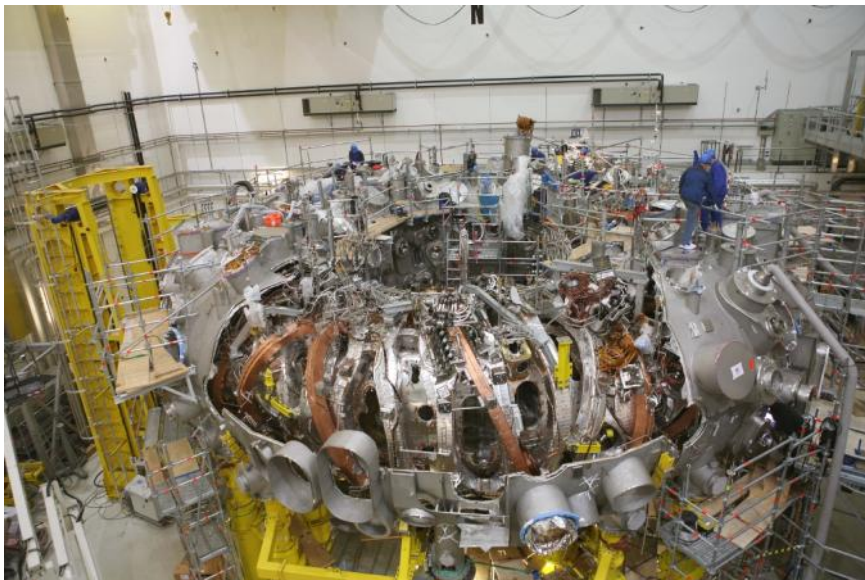
Not shown in fig. 1 are the vacuum pumps, the vessel ports, the electric cable feedthroughs, the water cooling pipes, the current bus-bars, the superconductor current joint system. One has to keep in mind that the cryostat vessel is large (volume about  $420 \text{ m}^3$ ) and contains the superconducting coil with a weight of about 5 tons each, houses a very complex superconducting bus-bar system, many kilometers of different types of helium and water supply lines with thousands of welding seams, many kilometers of sensor cables and hundreds of electric cable feedthroughs.

Pressure and temperature gauges measure the helium pressure  $p$  and temperature  $T$  at the two inlet pipes, labeled by “M” and “K”, and the outlet pipes labeled by “L2N2”, “L1” and “N1”. Each coil is equipped with up to eight additional temperature gauges. No further pressure gauges or flow meters are available inside the cryostat. The helium leak detectors provide the helium leak rate  $q$ , measured for the leaks between the helium pipes, the cryostat air side and the cryostat inner vessel volume. Due to the absolute calibration of the leak detectors, the leak rates are available as a  $pV$ -flux, i.e. in units of mbar l/s. The major part of the leaks had been detected and fixed during the standard leak search campaign in the year 2014. The remaining leaks exist between the helium pipes and the cryostat vacuum. Their integral size is in the order of  $q \approx 3 \cdot 10^{-5}$  mbar l/s at 4 bar helium pressure. In addition, leaks remain between cryostat and air with an integral leak rate of  $\approx 4.6 - 9.6 \cdot 10^{-4}$  mbar l/s. Just as the leaks

between helium pipes and cryostat, they turned out as too small to interfere with the coil operation. Due to the huge cryogenic pumping power of the cold surfaces, the air is immediately adsorbed on the cold surfaces and does not disturb the machine operation. Finally, the helium pipes have leaks towards the air, their leak size is unknown.



*Fig. 2. Photograph of the installations inside the W7 cryostat. One can see a small part of the vacuum vessel, the nonplanar and planar superconducting coils, the conductor bus-bars and the helium tubes, the water tubes and the bundles of the sensor cables.*



*Fig. 3. Photograph of the entire torus, just a few days before the cryostat vessel was closed. One fifth of the cryostat skin remains to be closed at this stage. One can see the entire torus with the port openings, reaching through into the cryostat and plasma vessel and the nonplanar and planar superconducting coils. Still missing are some of the port tubes, some*

*mechanical supports and the multi-foil super-insulation. The technicians give an impression on the size of the experiment.*

The photographs in the figs. 2 and 3 show a part of the interior of the cryostat. Besides all installations visible, some of the mechanical supports are still missing, as well as the super-insulating multi-foil layers. Those were of particular concern, because they provide a large surface area inside the vessel, which can adsorb water and gases. In addition, they reduce the guidance values inside the cryostat vessel for the gas flow during pumping down, and for the leaking helium from the leaks on its way to the pumping stations. To minimize the risk of leaks, all tube connections are welded.

Within the measurement time interval 14.02.2015 and 8.03.2015, the W7-X cryostat was cooled down from room temperature of about 290 K to the working temperature of about 4 K. For our numerical investigations, the slightly longer time interval between 4.02.2015 and 26.03.2015 is considered. This provides a few days of constant starting (room) temperature at the beginning, and at cryogenic conditions at the end of the measurement time interval. In the following, emphasis is on the measurement of the time traces  $p(t)$ ,  $T(t)$  at the five pipes “M”, “K”, “L2N2”, “L1” and “N1”. The leak detector signal  $q(t)$  was taken on one single pumping station with one and the same leak detector, measured continuously throughout the entire time interval. However, the measurements were not completely undisturbed, because several technical working teams were busy on W7-X doing technical tests, system commissioning, last welding activities, water temperature excursions on the cooling water pipes etc. Further helium leak search activities in vicinity of the torus caused, from time to time, a temporal increase of the measured helium flux  $q(t)$  through the remnant leaks towards the air. Therefore the measured helium fluxes  $q(t)$  will contain a certain amount of noise, i.e. variations of  $q(t)$  without a clear correlation to  $p(t)$  and  $T(t)$ . Another noise source is the leak detector itself. During the long measurement period the detector sensitivity can be subject to fluctuation caused by electronic noise, temperature variations, or a drifting of the detector working point. A constant helium temperature is assumed for the detected gas, because the helium has to flow a distance between a minimum of  $\approx 18$  m and a maximum of more than  $\approx 50$  m inside the vessels and ports to the leak detector. In addition, it has to cross the turbo pump, while all surfaces are at ambient temperature. Furthermore, the localizations of the leak(s) does not change in time. Therefore we assume that the leak detector efficiency is not affected by helium temperature variations.

A first localization of the leaks between the helium pipes and the cryostat was performed during the standard leak search in 2014. It provided preliminary results, for instance using a run-time method [5]. But for the later repair in the future a much better localization is desired. This localization and characterization is of crucial interest, as the cryostat allows only for very limited human access. The conditions are complicated by the fact that all hardware components are densely packed inside the vessel, and that the number of service ports is small. Very often large spatial distances exist between these ports and the helium filled components. Finally, the number of potential positions under leak suspect is high. This future leak search can later be done probably only by the helium sniffer method under air. This is a rather slow and not very precise method, taking the complicated space situation inside the cryostat into account. The leak rate in the range of  $q \approx 3 \cdot 10^{-5}$  mbar l/s seems to be easily accessible for the sniffer method, however many potential leaks are hidden deeply underneath hardware components, all surfaces are heavily polluted with helium, the area of surfaces is large because of the multi-layer super-insulation, and if the leak(s) are localized remains the task of fixing it by re-welding, under the difficult conditions described. Any information prior to this future leak search inside the cryostat will therefore be extremely helpful and might save time and man-power.

One has also to consider that each warm leak bears the risk of widening more under cryogenic conditions, such that the hazard of a larger leak rate in the future is given. This makes the repair of the remaining leaks even more important.

In this paper, a numerical regression analysis is described that is employed to localize the helium leak(s) better. A characterization of the helium flow regime inside the leak channel(s) is attempted. The numerical techniques described in the following are hence a supplement to the standard methods that had been used before. The available measurement data are exploited to a higher level than this was done before, and the numerical regression could complement the standard procedures in case of doubt or ambiguity.

## 2. Definition of the used quantities

In this chapter, the relevant helium parameters and their role for the regression analysis are described. The regression analysis and the regression models are briefly described, as well as some statistical definitions. In order to compare the different regression results to each other, the assessment of the quality is done by the Coefficient of Determination  $R^2$ . The numerical calculations are performed by FORTRAN routines and MATLAB scripts.

The laminar viscous gas flow in a leak channel is referred to as Poiseuille flow in textbooks [7] according to the formula  $q = \frac{\pi r^4}{16\eta l} \cdot (p_1^2 - p_2^2)$ . The molecular gas flow at lower pressure is defined as Knudsen flow according to  $q = \frac{\sqrt{2\pi}}{6} \cdot \sqrt{\frac{RT}{M}} \cdot \frac{d^3}{l} \cdot (p_1 - p_2)$ . The pressure transition regime in between these two, according to Burrows [8], is defined as the sum of a Poiseuille and a Knudsen flow formula. In this transition regime, the flow velocity reaches sound velocity inside the leak channel and blocking occurs. Turbulent viscous flow follows the relation:

$$q = d \left\{ \frac{20\pi^2}{51,2} \cdot \frac{d^3 \cdot (p_1^2 - p_2^2)}{2l} \right\}^{\frac{4}{7}} \cdot \left( \frac{RT}{M_{mol}} \right)^{\frac{3}{7}} \cdot \left( \frac{4}{\pi \cdot \eta} \right)^{\frac{1}{7}} \quad (1)$$

Note the dependency  $q \propto p^2$  for the laminar viscous flow,  $q \propto p$  for the molecular flow and  $q \propto p^{\frac{8}{7}}$  for the turbulent viscous flow. In these formulas,  $d$  stands for the leak channel diameter and  $r$  for the radius,  $l$  for the leak channel length,  $\eta$  for the dynamic viscosity,  $R$  for the molar gas constant,  $M_{mol}$  for the molar mass,  $p_1$  for the high gas pressure (inside the pipes) and  $p_2$  for the low gas pressure (vacuum) side. As the precise shape of the leak channel is unknown, the definition of a constant leak channel diameter  $d$  or radius  $r$  is of course not possible, i.e. that concept is strongly over-simplified. However, the numerical investigations will refer only to correlations between  $q(t)$  and analytical functions of  $p(t)$ ,  $T(t)$  and  $l/\eta(t)$ . This holds regardless of the shape of average diameter of the leak channel. Therefore this aspect plays no role within the scope of the following calculations. It has to be mentioned, that during the last years the numerical treatment of stationary gas flows through channels could successfully be evolved, providing new interpolation formulas across the different flow regimes. Details and more references can be found in [7], chapter 5. Even for the case of varying leak channel cross sections, compact formulas are derived by F. Sharipov et al. for the calculation of the mass flow rate. In this context it is noted that the analytical dependency for the molecular flow regime is both  $q \propto p \cdot \sqrt{T}$  for the Knudsen and the Sharipov formula. In our case, the leak channel size, its length etc. are unknown, therefore we handle this lack of information within the regression analysis by the use of free regression coefficients (see below), i.e. for the assessment of the regression quality the constant of proportionality plays no role. Remarkable is the difference between the Poiseuille formula  $q \propto p^2/\eta$  and the Sharipov formula  $q \propto p^2/\eta \cdot T$  with the additional temperature dependence. For the

consideration here it is always assumed that  $p_1 = p \gg p_2 \approx 0$ , as typical for a leak flow into vacuum at  $p_2 \approx 0$ .

Within this paper, we will focus on the question: Can one localize the leak(s) within the closed cryostat? The idea is to use a regression analysis between  $p(t)$  in the five pipes on the one hand, and the helium-flux  $q(t)$  measured by the leak detector on the other. The flux  $q(t)$  will be a function of the pressure, temperature and dynamic viscosity of the helium. Any temporal pressure, temperature or viscosity variation will cause a variation in  $q(t)$ . The stronger the correlation between functions of  $p(t), T(t), \eta(t)$  on the one side, and  $q(t)$  on the other, the closer should the leak(s) be to the pipe under consideration.

As one is free to choose a regression model according to a linear (molecular flow) and a quadratic (viscous flow), numerical attempts were made for a variation of the regression model. The motivation was the question, whether the type of model with the best (or better) regression can tell us something about the flow nature in the leak channel.

During the measurement time interval, the helium leak tester values for  $q(t)$  are recorded almost permanently. Due to the leaks between the helium pipes and the cryostat, always a finite helium-flux is observed. In parallel, the inlet and outlet temperature and pressure of the helium in the five pipes mentioned above are measured and stored. In order to obtain a common time basis for the sampled measurement points for all signals, a numerical re-sampling is done by a MATLAB [9] routine using an interpolation algorithm. For that step, a linear interpolation between the individual data points of one time vector is done at the points in time of the other time vectors. As the sampling rates for the measurement of  $p(t), T(t)$  and  $q(t)$  are the highest, these time bases are used as reference for the other signals. Thus, no data points are lost.

The common time-basis is finally expressed in units of ns in UTC (Coordinated Universal Time). Thus, several millions of sampled data points are obtained for  $p(t), T(t)$  and  $q(t)$ . In order to calculate numerically the correlation between  $q(t)$  on the one side and  $p(t), T(t), \eta(t)$  on the other, a MATLAB routine is written that performs a numerical regression between these parameters. The analytical type of regression function can be chosen freely. To take into account the supercritical helium dynamic viscosity  $\eta$  correctly, the FORTRAN routine “*He\_prop*” [10] is used. It calculates, among other helium properties,  $\rho$  and  $\eta$  from  $p$  and  $T$ . Finally, numerous regression calculations are performed between  $q(t)$  on the one hand, and, on the other for the following types of regression:

- a)  $\alpha + \beta \cdot \frac{p^2}{\eta}$  for the check for a possible Poiseuille flow behavior



- b)  $\alpha \cdot p \cdot \sqrt{T}$  for the check for a possible Knudsen/Sharipov flow behavior
- c)  $\alpha + \beta \cdot \frac{p^2}{\eta} + \gamma \cdot p \cdot \sqrt{T}$  for the check of a possible Burrows flow behavior  $q(p)$
- d)  $\alpha + \beta \cdot \frac{p^{8/7}}{\eta^{1/7}} \cdot T^{3/7}$  for the check of a possible turbulent viscous flow behavior

always as a function of the time.

The parameters  $\alpha$ ,  $\beta$  and  $\gamma$  are the free regression coefficients. They are calculated for each regression analysis anew.

Within the available MATLAB subroutines, the quality of the regression is calculated using the quantity  $SSE = \sum_{i=1}^n w_i \cdot (y_i - \hat{y}_i)^2$ , the Sum of Squares due to the Error, the quantity  $SSR = \sum_{i=1}^n w_i \cdot (\hat{y}_i - \bar{y})^2$  the Sum of Squares of the Regression, and the Total Sum of Squares =  $SST = \sum_{i=1}^n w_i \cdot (y_i - \bar{y})^2$ . These definitions hold for a number  $n$  of data points  $y_i$  in the regression with the weight factors  $w_i$  and the average value  $\bar{y}$ . The  $\hat{y}_i$  are the regression values. A convenient parameter for the assessment of the quality of the regression is the Coefficient of Determination  $R_{squares}$ , defined by  $R_{adj}^2 = 1 - \left(\frac{n-1}{n-m}\right) \cdot \frac{SSE}{SST}$  in MATLAB. It is adjusted to the number  $m$  of free regression coefficients that is, for the relatively simple numerical problem and assumptions presented here, always between 1 and 3. The resulting value of the adjusted  $R^2$  is equal to 1 for a perfectly fitting regression model and = 0 for no correlation at all. The use of the adjusted  $R^2$  allows for the comparison of different types of regression, even for a varying number of regression coefficients.

In order to avoid the comparison of completely different phases of the cooling-down (see below), the entire measurement time interval is divided into shorter time windows for some of the calculations. In particular the time window with cryogenic conditions ( $T < 20K$ ) is examined separately, because here the helium properties and the behavior of all cryostat components might deviate strongly. For some regression calculations, however, the entire measurement time interval is used. In addition, the helium density and dynamic viscosity are pre-calculated and temporarily stored in tables, prior to the regression analysis runs. For the regression itself, these values are then fetched after interpolation from these tables.

### 3. Results and discussion

In this chapter, some of the results of the numerical regression analysis are presented. Fig. 4 shows, as an overview, the time traces of  $p$  and  $T$  during cool-down of the cryostat. Those data are measured at the inlet of the pipe labeled “L2N2” and show the variation of the helium

inlet parameters, as provided by the cryogenic facility. These are some of the raw data that are used for the later regression analysis.

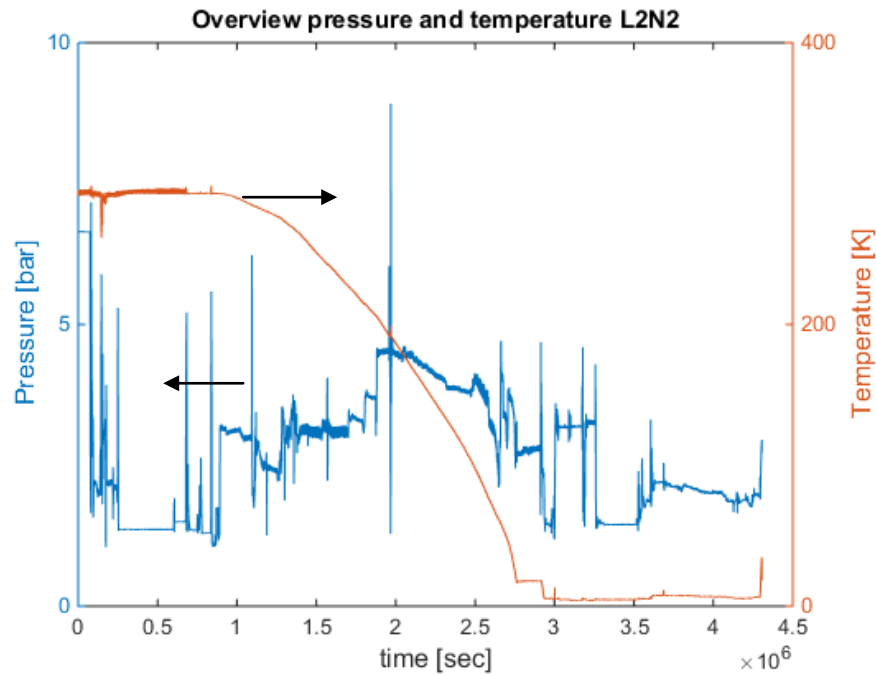


Fig 4: Time traces for the helium inlet pressure and temperature measured at the inlet pipe “L2N2”. The time axis starts with  $t=0$  s at the beginning of the measurement campaign.

Note that only at the end of the cooling-down, i.e. the last quarter of the period shown, cryogenic helium temperatures  $< 20$  K are reached. The most important question is about the localization of the leaks. As the values  $p$  and  $T$  for the helium flow are measured at 5 inlet/outlet pipes, only these 5 locations are accessible to the regression analysis. Figs. 5 and 6 show the results for the  $SSE$ -values and adjusted  $R^2$  for the five pipes mentioned. Shown here are only the results for two types of regression (linear and quadratic) for the check whether viscous or molecular flow dominates the leak flow. We can restrict the presentation of the results on these two regression types, because they already show the desired finding. The  $SSE$  and adjusted  $R^2$  were calculated as the average of all time windows which fulfilled the condition for the adjusted  $R^2 > 0.1$ .

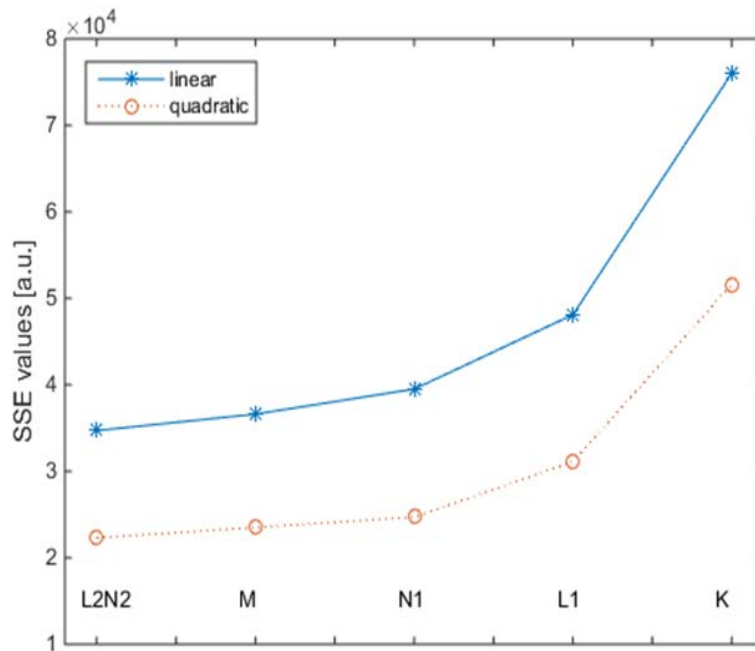


Fig 5: SSE-values calculated for the 5 pipe systems labeled L2N2 up to K. The pipe labels are indicated on the plot. The solid line (stars) shows the results for the linear model b), the dotted line (circles) for the quadratic model a).

Fig. 5 shows, that the values for  $SSE$  are smallest for the pipe “L2N2” and largest for “K”. This indicates that the best correlation is given at pipe “L2N2”. This is consistent to the results for the adjusted  $R^2$  which are largest (the highest regression quality) for “L2N2” and smallest for “K”, see fig. 6. All other values for  $SSE$  and the adjusted  $R^2$  lie well ordered in between these two extreme values, almost always in the same order. Only the values for the pipes “L1” and “N1” are reversed when comparing  $SSE$  and adjusted  $R^2$ , possibly indicating that the regression for these two pipes is more or less of the same quality.

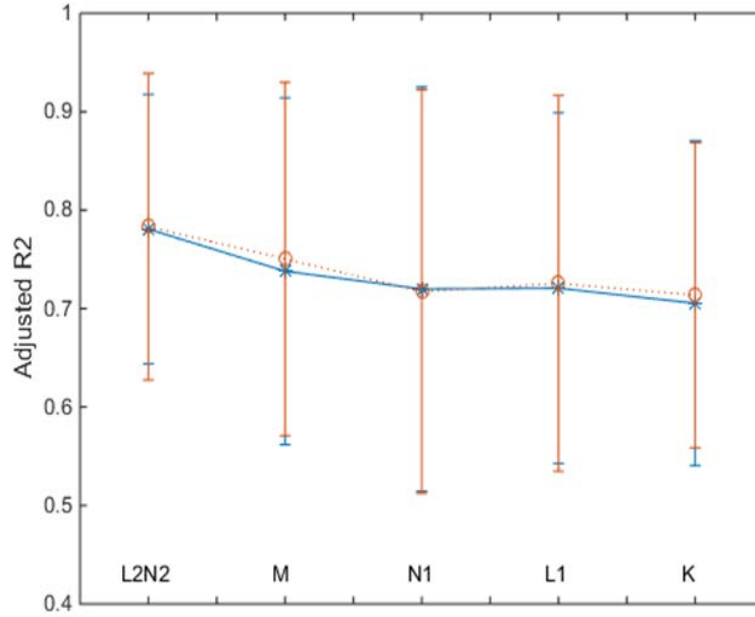


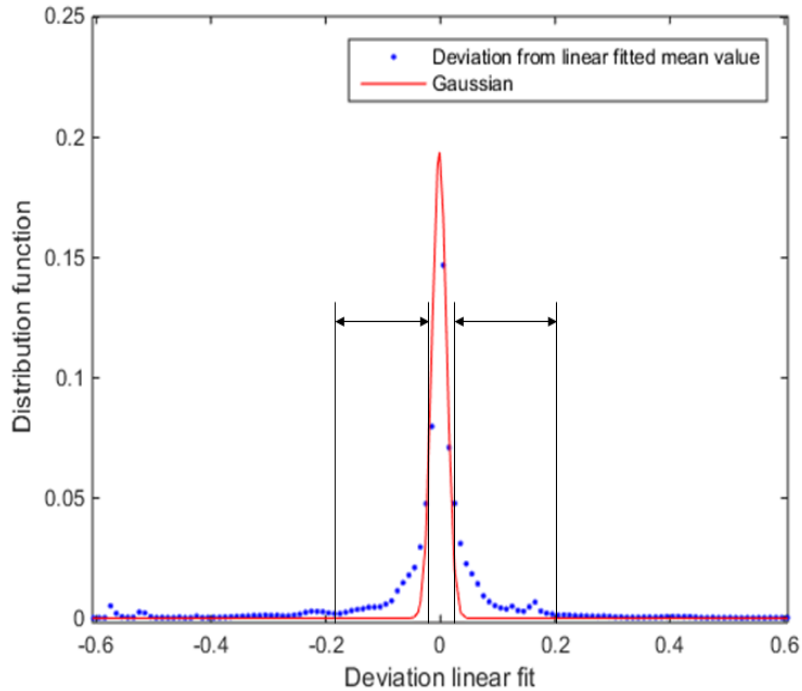
Fig 6: Results of the calculated adjusted coefficients  $R^2$  for the 5 pipe systems labeled L2N2 up to K. The pipe labels are indicated on the plot. The solid line (stars) shows the results for the linear model b), the dotted line (circles) for the quadratic model a). The error bars show the size of one standard deviation.

One has to keep in mind that all pipes are connected to each other, however with different hydraulic resistances. Therefore the regression results for all positions are very close to each other, but the ordering in leak size determines the quality of the regression. Hence one cannot tell whether, for instance, the pipe “K” is completely free of leaks, as a certain correlation is given between  $q$  on the one hand and  $p$ ,  $T$  and  $\eta$  on the other. These hydraulic connections are a draw-back for the regression analysis, however when comparing the size of the regression quality one can tell in which pipe the total leak rate is largest, and in which lowest. The findings described here are consistent to the results that had been obtained during the standard leak search [5]. There, also the pipe “L2N2” was the one with the strongest indication for a leak.

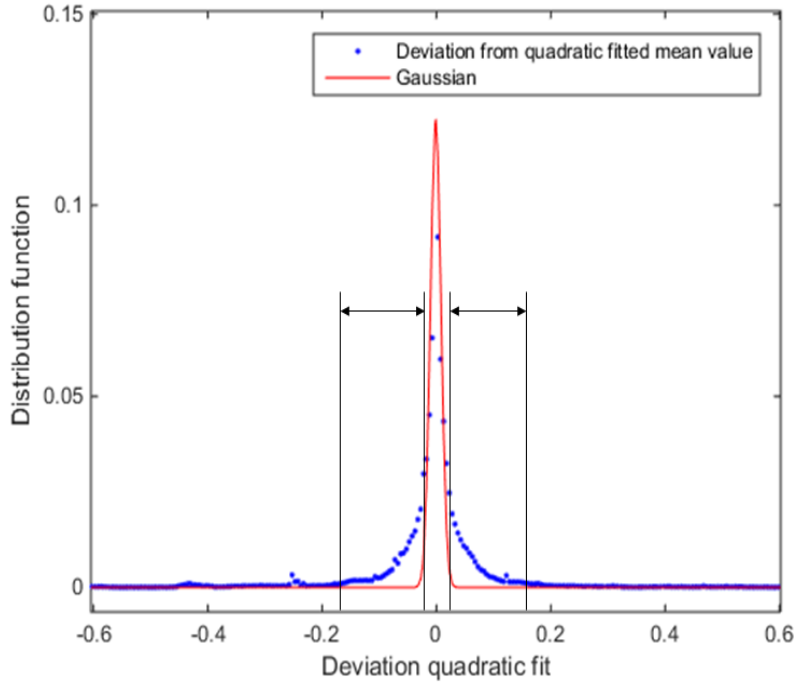
From the figs. 5 and 6 alone, one cannot distinguish which type of regression (linear or quadratic) fits better. Nevertheless, the two figures show that all quadratic regression models type a) fit better than the linear regression models type b). All errors  $SSE$  for type a) are smaller than for type b) in fig. 5, and the adjusted coefficient  $R^2$  for type a) are always slightly larger than for type b) in fig. 6.

Fig. 6 shows also the calculated values of one standard deviation  $\sigma$ , indicated as error bars. They were calculated from the deviation between the regression result and the measured

values, independently for the five pipe systems. Obviously, the slight difference in the adjusted coefficients  $R^2$  between the five pipe systems is statistically not significant, because it is always smaller than all standard deviations. However, this statement has to be taken with some care, because the statistical significance is a question of the probability distribution function. The use of  $\sigma$  assumes a Gaussian process. The statistical significance of  $\sigma$  for our investigations is now investigated briefly. The figs. 7 and 8 show plots of the calculated distribution functions  $F = P(y_i - \hat{y}_i)$  with  $P$  being the probability that the deviation  $(y_i - \hat{y}_i)$  is realized. For these plots, the pressure measurements in the pipe system L2N2 are used. The integral of the distribution functions is normalized to one. For the quantitative comparison between the Gaussian fit and  $F$ , the Full-Width Half-Maximum FWHM of the Gaussian (to 95% confidence) and  $\sigma$  of  $F$  are calculated.



*Fig. 7: Plot of the distribution function of the deviation between regression and measured pressure values in pipe system L2N2. The dots show the distribution function for the linear regression, the solid curve shows the fit to a Gaussian. The curves are normalized:  $\int F = 1$ . The horizontal arrows mark the region where the calculated distribution function deviates remarkably from the Gaussian fit.*

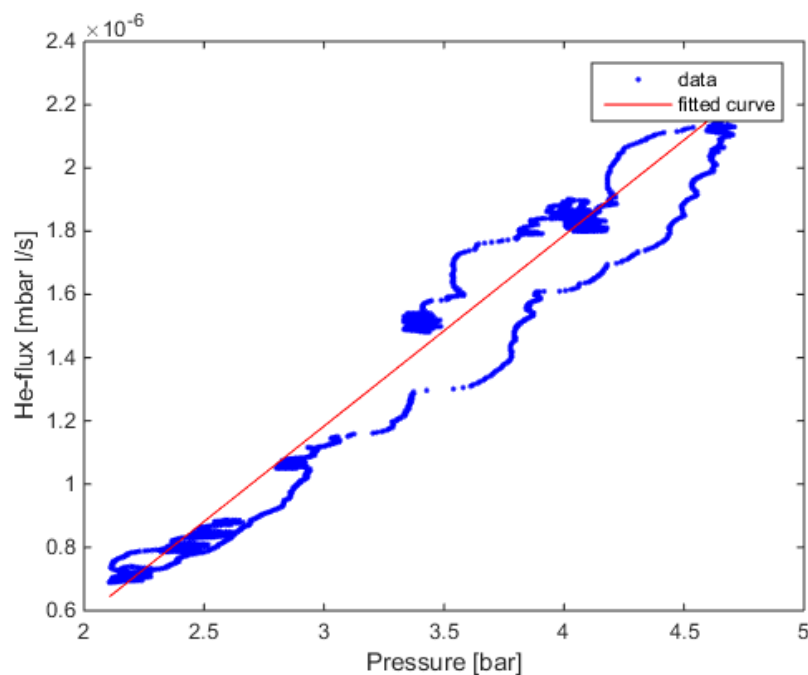


*Fig. 8: Plot of the distribution function of the deviation between regression and measured pressure values in pipe system L2N2. The dots show the distribution function for the quadratic regression, the solid curve shows the fit to a Gaussian. The curves are normalized:  $\int F = 1$ . The horizontal arrows mark the region where the calculated distribution function deviates remarkably from the Gaussian fit.*

As can be seen from the figs. 7 and 8, the Gaussian curves fit not well to the calculated distribution functions  $F$ . This means, that obviously the deviations between regression and measurement are not only induced by statistical error sources, like detector noise etc., but also to a considerable amount by systematic influence from outside. This was already mentioned above. Therefore,  $\sigma$  alone is no good measure to assess the statistical significance, as this was tried for the data shown in fig. 6. For fig. 7, the Gaussian FWHM =  $\pm 0,0178$  and  $\sigma = 0,22$ . For fig. 8, the Gaussian FWHM =  $\pm 0,013$  and  $\sigma = 0,18$ . Hence, the Gaussian FWHM is by a factor of  $\approx 10$  smaller than  $\sigma$ . Nevertheless, the use of  $\sigma$  reflects the impact of all perturbing effects and is considered here as relevant, however not in the strict statistical sense for the distribution function as found for the measurement data and the regression analysis.

Therefore, we can only draw the conclusion that the pipe system around L2N2 might have the highest probability for a leak and start the leak search there. The final confirmation for the positions of the leak will come, when the cryostat will be opened again, the leaks will be searched by conventional methods, and can eventually be fixed. But for the next months to come, plasma operation will have the first priority, not the leak search.

The additional question addresses to the type of helium flow inside the leak channels. To tackle this task, the different types of regression a) to d) were tested for the five different pipe locations. The fig. 9 shows one selected example of the regression analysis. It is taken from the pipe labeled “L2N2” because here the regression quality is the best. It is one of the best results (and therefore not typical) obtained for a linear regression type b) for the check for a possible Knudsen flow inside the leak channel, plotted as the measured data points  $p(t)$  versus  $q(t)$ . The plot shows the measurement data points as dots, as well as the result of the regression. To some extent, the regression represents the measurement points, but obviously a further, unknown parameter provokes a systematic deviation.



*Fig 9: Plot of the regression according to a linear model, plotted as  $q(t)$  versus  $p(t)$ . Shown are the individual measured data points (dots) together with the result of the regression analysis (solid line). Adjusted  $R^2 = 0.977\dots$  in this case. Obviously, the inlet pressure makes an excursion between about 2 bar and 5 bar within this time window, and the helium leak flux reacts to the pressure change.*

To make it short: unfortunately the comparison of the different regression models provided no conclusive results. In particular during the time period with  $T < 20$  K, too many other disturbing effects blurred the correlation between  $p(t)$  and  $q(t)$ .

In the following, the results are briefly discussed. With the definition for the Reynolds number  $Re = \frac{4M_{molar}}{\pi\eta RT} \left( \frac{q_p V}{d} \right)$  one finds Reynold numbers  $< 1$ , and therefore  $Re \ll 2300$ , hence a value which is far below the threshold for the transition between laminar and turbulent flow.

Here,  $M_{molar}$  is the molar mass value,  $R$  the molar gas constant,  $q_{pV}$  is the  $pV$ -leak flux. Thus one can exclude the turbulent viscous flow. That conclusion is confirmed by the observation that the regression analyses never showed clearly, whether the dynamic viscosity plays also a decisive role or not. The strongest impact on the quality of regression came always from the pressure dependence. This indicates a dominating molecular flow.

Obviously, other parameters play a role for the helium leak flux than only a change in helium pressure or dynamic viscosity. We suspect that possible “cold leaks” show the tendency to change their leak size with the temperature, maybe in a non-linear manner. Those must not be located at pipes with  $< 20\text{K}$ : inside the cryostat many other components exist which have higher temperatures. In fact there exist components at room temperature (water pipes, the outer cryostat skin), components at  $\approx 40\text{ K} - 60\text{ K}$  (the radiation shields), and inlet and outlet pipes with temperatures somewhere between room temperature and helium temperature. All those are coupled by radiation and rest gas heat conduction to each other in an unpredictable manner. For the possible increase of cold leaks see the quantitative estimates [11].

During the cooling-down of the cryostat it could be observed that helium leaks exist between the cryogenic valve box in the torus hall, and the air. Therefore, helium from these leaks will escape into the air. Through the air leaks in the cryostat vessel they will finally enter the leak detector. This effect frequently disturbed the standard leak search that was performed before, and obviously this effect plays again a disturbing role for the regression analysis. The same perturbations had also been observed, when helium leak search was performed at other components in the torus hall.

Furthermore, from time to time outbursts of gas inside the cryostat are observed, obviously spontaneous or without perceptible reason. Together with an increase in the helium flux, also heavier gases like  $\text{H}_2\text{O}$ , Ar,  $\text{N}_2$  or  $\text{CO}_2$  appear correlated with the helium flux. This was measured with the mass spectrometers, connected to the cryostat pumping stations. In that case, not only helium will produce a pressure increase but also the other gases. We speculate that a getter effect of helium on frost layers of heavier gases might play a role. Or the heavier gases are frozen out at warmer outlet pipes, which warm up simultaneously with the colder ones and release the heavy gases simultaneously with the helium. It can also not be excluded that helium will temporarily be adsorbed on metallic surfaces if temperatures around  $4\text{ K}$  are reached, and can be released later. This all will strongly perturb the results of the regression analysis.

Finally, the ongoing technical tests and commissioning activities around the cryostat produced erratic local warming and mechanical vibrations of components with the subsequent release of



gas that will superimpose the helium leak rate. The impact of these perturbations was found to be the stronger, the colder the cryostat is.

#### 4. Conclusions

The ordering in size of the obtained  $SSE$  and adjusted  $R^2$  values provides, at first glance, an indication for the localization of the leak(s). The correlations are highest for the pipe labeled “L2N2”, followed by pipe “M”. This is consistent to findings from the standard leak search activities, where for instance helium run time methods were employed. Hence, the future leak search will concentrate on the vicinity of these pipes. Nevertheless, the final confirmation is pending until all leaks are localized and successfully closed. The assessment of the statistical significance of the regression analysis has to be done with some care, as the relevance of statistical parameters depends on the distribution function of the deviations from the regression. In case of doubt (as in our case) the regression analysis might only provide rough hints to the leak(s) position(s), but no proof.

Not conclusive are the results concerning the nature of the flow in the leak channel. The regression analysis is always perturbed by noise or by un-controllable experimental perturbations. Maybe on other machines with large cryostats the conditions are better or the correlation is stronger, providing more distinct results. Nonetheless the conclusion can be drawn that molecular flow conditions must be fulfilled inside the leak channel. The small size of the leak(s) will make the leak search more complicated, in particular because probably only the helium sniffer method can be employed, a difficult method for such small leak rates, and taking into account the boundary conditions in our cryostat.

Unfortunately, only five positions equipped with pressure sensors were available. Therefore this part of the search had to remain very coarse in space. With more positions for pressure gauges one could refine the numerical regression search for the location of the leaks, correspondingly. The search could also be refined in space, if several leak detectors (instead of only one) were connected simultaneously to the five pumping stations around the W7-X torus. Finally, some closing valves within the network of helium pipes could also help for the localization of the leak(s).

## Acknowledgements

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

## References

- [1] Beidler C, Harmeyer E, Herrnegger F, Kisslinger J, Igitkhanov Y, Wobig H. Fusion Technol. 1990; 17: 148–68.
- [2] Grieger G, Lotz W, Merkel P, Nührenberg J, Sapper J, Wobig H, et al. Phys. Fluids B 1992; 4: 2081–91.
- [3] Klinger T, Baylard C, Beidler C, Boscary J, Bosch HS, Dinklage A, et al. Fusion Engineering and Design 2013; 88: 461-5.
- [4] Reich J, Gardebrecht W, Hein B, Missal B, Starke F, Seidler KU. Fusion Engineering and Design 2003; 69: 345-8.
- [5] Baldzuhn J, Reimer H, Biedermann C, Grote H, Hathiramani D, Kornejew P, Rademann D, Volzke O. Vacuum 2015; doi: 10.1016/j.vacuum.2015.02.014
- [6] Bosch HS et al. Fusion Engineering and Design 2015, article in press.
- [7] Jousten K: Wutz Handbuch Vakuumtechnik. 11th Edition, Springer Vieweg Verlag 2013.
- [8] <http://www.leakdetection-technology.com/>
- [9] MathWorks Inc., R2014b
- [10] Cryodata Inc.; Heprop (1999)

[11] Cruikshank P, Claudet S, Maan W, Mourier L, Perrier-Cornet A, Provot N,  
[accelconf.web.cern.ch/AccelConf/IPAC2011/papers/tups020.pdf](http://accelconf.web.cern.ch/AccelConf/IPAC2011/papers/tups020.pdf)

Reviewers' comments:

Reviewer #1: I have enjoyed a lot reading the manuscript, which in my opinion deserves publication.

The mathematical technique used is rather interesting, and could be applied to similar problems in large accelerator or fusion machine systems (CERN, ITER, for instance).

The paper is clear, it is well written, and gives sufficient references to the reviewer.

The only thing I would add is a reference to fig.2 ("Density variation of supercritical helium vs temperature"), or a statement "measured by authors" in case it has not been taken from literature.

Has been included.

On page 14, the authors write: "We suspect that possible 'cold leaks' show the tendency to change their leak size with temperature, maybe in a non-linear manner.", and I can only point them to the paper "Leak Tightness of LHC Cold Vacuum Systems", Proc. IPAC-2011 Conference, San Sebastian, Spain, where they claim that...

"In cases where helium leaks had been measured prior to cool down, the equilibrium pressure at saturation can be used to estimate the increase in leak magnitude under cold conditions. The limited data give values between 100 and 7000. "

(Paper here: [accelconf.web.cern.ch/AccelConf/IPAC2011/papers/tups020.pdf](http://accelconf.web.cern.ch/AccelConf/IPAC2011/papers/tups020.pdf) )

Thanks for that hint, is included.

Reviewer #2: (see also attachment)

The paper deals with an interesting approach for leak search at complex devices.

Unfortunately no real results were obtained, which is not the problem.

But the author avoid a scientific discussion:

No errors are discussed, is there a significant difference for  $R^2$  0.7874 and 0.7364 ?  $R$  values of up to 6 digits are given (fig 5 ).

A good remark. In fact, I had doubts on how to define an error of  $R^2$ , because  $R^2$  itself already defines something like an error (or better, a normalized deviation between two data-sets). The 6 digits are clearly too much, I fully agree. The question is then about the statistical significance of the coefficient of determination. The answer depends on the type of distribution function of the "scattering" distribution around the regression. This distribution function is unknown. See more below at (\*).

The statistical approach seems to be quite simple: ' despite the fact that many matlab regressions calculations were performed'

A more sophisticated model of the system should be used for statistical analysis.

For the models the flow characteristics were chosen, as defined for the different flow regimes. Which more sophisticated model could that be, except those pre-determined by the flow physics?

The authors just pick 6 different models for fitting and try to find correlations.

And I was afraid that already 6 might be too much / too puzzling?

In Fig 4 a linear one dimensional fit was shown, which may cover the general tendency, but the data depend not only on pressure, so multi dimensional fits are needed.. No other fits are shown for comparison.

As a timesequence, for each point in time we have a dedicated set of  $p_i$   $T_i$  ,  $\eta_i$  ,  $\rho_i$  etc, no multivariate data set. The different models cover exactly, what one multi-dimensional fit will do. However, the models are motivated by the different flow regimes where we want to know which of them is decisive for the leak flow. Therefore, squeezing all variables into one model will not help with respect to that. The opposite is true: they have to be treated separately with respect to the different flow regimes. Only then we could, at least in principle, separate between a linear or quadratic behaviour.

However, it failed unfortunately. A remark is included in the text (in 4. Conclusions), indicating that this method can therefore only be considered as a suggestion for other machines. Maybe on the other machines the conditions are better for such an analysis, providing clearer results, or the other authors find a better flow model for their description. In this case, even my negative outcome might help them further.

In Fig 6 the results for fitting an exponent are shown. A smooth curve without any noise is presented !

The variation of  $R^2$  is in the order of  $(0.23-95-0.2365)/0.238 = 1\%$ . I wonder to get this accuracy for  $r^2$ .

(\*) To answer this question, more calculations are included in the text. There, the accuracy and the statistical significance of the results is discussed in more detail, and new figures are included showing the relevant distribution functions and their consequences for the assessment of the results of the regression analysis.

In the real world there are many disturbances of the He signal due to other activities. The author mentioned due to this fact they have to bin the data. But no information on the bins and the selection criteria are given.

A remark is included. In fact, no binning in the sense of an averaging was done.

Comments are also given in the pdf file enclosed.

See the corresponding answers in your comment windows and the manuscript.

Reviewer #3: (1) The proposed method is interesting and expected to give useful information for the localization of the leak. The obtained results, however, are not conclusive yet. It seems that some modification is needed for the data analysis. For example, have you considered the effect of the helium temperature on the leak detector? Since the conductance of molecular gas flow is proportional to the gas velocity, the gas temperature might affect the sensitivity of the leak detector.

This is correct. However, the manufacturer of the leak detectors provides no calibration curve versus temperature. A remark is included in the text. The helium travelled between 18 m and 50 m, the last 18 m doing collisions with the port wall under ambient temperature, and after crossing the turbo pump. Therefore, we assume a constant helium temperature. Furthermore, the localization of the leak does not change in time.

Therefore, I recommend that this paper is submitted again with some improvement after the leak is fixed.

The next chance to fix the leaks might come in 1-2 years, maybe even later. Maybe, we can never localize and fix the leaks; this is also possible. Waiting for that seems to be not an option.

(2) Minor comments are as follows:

a) P5, line 24: The equation for turbulent viscous flow must be referred.

Is done.

b) P7, line 5: "a numerical resampling" should be explained briefly, because the reference 9 is not easy to access.

Is done.

c) P8, line 13: " $\gamma_i$ " should be explained.

Is done.

Reviewer #4: The authors report about their efforts to commission the very large vacuum system of the plasma chamber and cryostat of the W7-X experiment. The leak localization of such a complex vacuum system is a huge challenge. The paper reports about an approach that tries to do the best for a system that is not designed to ease vacuum leak search.

The authors are defining three tasks, of which task 1 is clearly the most important, namely to demonstrate that the system which provides a good regression analysis between  $q$  and  $p$  is the one likely to be leaky. The authors conclude that this task has been managed well, but this is not (yet) the case. They found that the quality of the regression analysis varies from pipe system to pipe system, but their argumentation would only be confirmed if they would have found the leak really on the pipe system with the best regression fit. However, this last, but important step has not been done yet.

I fully agree. However, the next opportunity for leak search will come maybe in 1-2 years, but without guarantee that we will find the leaks. And if we can localize them, there is no guarantee that we can fix them (technically).

Waiting so long, with the risk of no success in the sense that you mention, seemed to be no option. The corresponding parts of the text are rephrased.

Task 2 about the flow type and leak size did not provide convincing results and I also do not share that this task is important. Neither is task 3 on the deviation between real gas and ideal gas behavior. If the thermodynamic tables are implemented correctly, and this is what the authors claim, I do not see this to be an issue anymore.

The tasks 2 and 3 are deleted from the paper, i.e. are only briefly mentioned, and that no conclusive results were obtained.

In consequence, I suggest that the authors revise this paper significantly and focus on task 1 only. Only then, and although I am still not particularly convinced about the statistical significance of the experimental findings, I think, the paper is worthwhile to be published.

Is done. In particular, concerning the statistical significance some new data are included, and the role of the distribution function is discussed briefly.

Furthermore, the following minor points should be addressed in the revision:

1. I personally do not like references to be included in the abstract (but the journal editor may think differently). If doable, I encourage to remove them from the abstract and add them to the main text.

Is done.

2. The text is difficult to read and understand when it comes to the description of the system complexity. Here, to add more figures (piping diagrams) or photos may be helpful.

Two photographs are included, plus some description in the text.

3. One reason for the challenge associated with this vacuum leak task comes from the fact that there are not sufficient valves for segregation of parts of the flow network. Of course, on the other side, installing more valves adds also complexity and increases the failure rate. I would be interested to hear a discussion on this

aspect and a recommendation for future machines such as ITER from the view of the vacuum experts in charge of the leak search.

A short discussion is included.

4. It is not state-of-the-art anymore to treat the transitional flow as weighted sum of free molecular and laminar flow (what the authors call 'according to Burrows'). Correspondingly, the reference [8] of 1982 is superseded, at least the three latest editions include a dedicated chapter on vacuum gas dynamics with a full treatment on transitional flow. Within the authors' simplified approach, this may not be a problem, but for educative purposes, I ask the authors to add an additional section on this issue. There is also a bunch of literature data existing on modelling flows through unknown leak geometries.

Thanks for this valuable hint. A corresponding brief discussion is included in the text.

5. Fig. 2 and the whole discussion on ideal vs real gas properties is not needed to be discussed in this detail, as the authors are using HEPROP anyway. I strongly suggest to delete Fig. 2 and the associated discussion.

Is deleted, as well as the discussion about the real gas properties.

6. Fig. 3 needs re-work in terms of the x-axis (to use numbers in the dimension  $10^9$  and the units ns is not good, why not go for s directly). Is now in sec. (Already before it was in sec, but the caption "nsec" was wrong).