**Detailed description of Ar-Ar results.**

**IR22:** Two incremental heating experiments were performed. A1\_1 does not yield a plateau according to generally accepted criteria (>50% 39ArK released in plateau; at least 3 consecutive steps overlapping at 2 sigma). We observe however an age of 49.99 ± 1.07 Ma (step 14-18; 33.5% 39ArK) for initial heating steps with higher K/Ca ratios (~0.194) and an age of 35.98 ± 1.11 Ma (step 24-34; 35.1% 39ArK) for higher heating steps with lower K/Ca ratios (~0.057). The total fusion age arrives at 44.00 ± 0.50 Ma. The 40Ar/36Ar ratios obtained from inverse isochrons do not deviate from atmospheric argon in both cases. Note, that measured intensities for the samples are rather low: <10 times blank on 40Ar. A1\_2 yields higher intensities for the individual steps and shows similar behaviour as A1\_1 with older ages around 52 Ma and higher K/Ca ratios during initial heating steps and younger ages around 37 Ma and lower K/Ca ratio during higher heating steps. The total fusion age of 43.73 ± 0.48 Ma overlaps with A1\_1. Radiogenic 40Ar content is rather low with the majority of the steps between 30-60%. Due to the heterogeneous age spectrum, it is difficult to obtain a reliable 40Ar/39Ar age for this sample. This also implies caution for the use of any K-Ar dates of these groundmass samples.

**IR216:** We performed two incremental heating experiments on groundmass and two on plagioclase from flow IR216. Both experiments show the same trend in K/Ca and age spectra are well defined. A5\_1 consists of only 6 heating steps and A5\_2 of 10. The two plateau ages are slightly offset (40.40 ± 0.13 Ma and 39.71 ± 0.09Ma). Isochrons do not show deviations from atmospheric argon and radiogenic 40Ar\* contents are high, generally >90%. The groundmass samples are slightly older (again two slightly offset plateau ages of 41.28 ± 0.10 Ma and 40.83 ± 0.11 Ma) with decreasing age spectra, high radiogenic 40Ar\* contents (mostly >97%) and high K/Ca ratios.

**IR215:** We performed two fusion analyses on single feldspar grains of this flow. Due to technical problems (and when these were fixed >1 year after irradiation) we were not able to obtain more data on this sample. Because of the low sample intensities (3-5 times blanks) the analytical uncertainties are high. The weighted mean age yields 39.34 ± 1.08 Ma and the radiogenic 40Ar\* content is >80%.

**IR119:** We performed two incremental heating experiments on groundmass from flow IR119. Both experiments show the same trend in K/Ca. Both age spectra show a slight decrease in age, with A2\_2 some higher ages in the final steps (cracking of inclusions?). The two plateau ages are slightly offset (39.71 ± 0.09 Ma and 39.46 ± 0.12Ma). Isochrons do not show deviations from atmospheric argon and radiogenic 40Ar\* contents are high, generally >97%.

**AZ16A:** We performed four incremental heating experiments on groundmass and three on plagioclase of flow AZ16A. Three groundmass samples show similar K-Ca trends, while the other shows higher K-Ca ratios during the first steps. Three groundmass isochrons show no deviation from atmospheric argon, age spectra overlap. Radiogenic 40Ar\* contents are high, above 84%. Two of the three plagioclase samples show overlapping age spectra. The third one shows lower intensities, and higher errors, with an age that is ~2 million years older. The deviating sample has a plateau of only two steps, and no isochrons can be calculated. Samples that show overlapping age spectra have similar K-Ca trends. One of the samples shows a slight deviation from atmospheric argon for the normal isochrons, all other isochrons do not deviate from atmospheric argon. Radiogenic 40Ar\* contents of samples with high intensities are >92%.

**AZ16M:** We performed four incremental heating experiments on groundmass of flow AZ16M. Results are variable, with three out of four samples showing high intensities. Only one sample shows generally high radiogenic 40Ar\* (>73%), but this sample shows a deviation from atmospheric argon on isochrons. The other three samples show low radiogenic argon, but no deviation from atmospheric argon on isochrons. K-Ca trends are variable.

**Si01:** We performed two fusion experiments on feldspar of ash Si01. In the first experiment, a small number of grains was fused in one step. Intensities are low, only 3-4 times above blank intensities. During the second experiment, a larger number of feldspar grains was fused, and higher intensities were obtained, resulting in a smaller error for this age. Both experiments give within error the same age. Radiogenic 40Ar\* is generally around 50%. Isochrons of the first experiment show no deviation from atmospheric argon, while those of the second experiment are slightly lower.

**X1.14B:** We performed three experiments on glass and two on biotite grains of ash X1.14B. The first experiment on glass grains gave low intensities, only just above blank levels. In the following experiment, more grains were fused, resulting in slightly higher intensities. Both experiments give within error the same age and isochrons of both experiments show no deviations from atmospheric argon. Radiogenic 40Ar\* contents are generally high; >77% with one exception (~55%). Three experiments on biotite result in much higher intensities compared to the experiments on glass. Radiogenic 40Ar\* contents are high, >89%. Two of three samples show isochrons that do not deviate from atmospheric argon. Age plateaus overlap, K-Ca trends are similar, and all three samples give within error the same age.

We add our new data to the compilation of literature ages, where we do not take into account our data of the same units as sampled by Vincent et al. (2005), as this would lead to oversampling of this unit and a bias in the age distribution. Of new units that were dated, we took ages of the ash layers that had the most measurements and smallest error; 41.28 ± 0.39 Ma for the Siyaki ash layer (Si01), and 44.78 ± 0.07 Ma for the Khilmili ash layer (X1.14B). For the igneous rocks in Iran, IR22 was left out, as this did not yield a reliable age. IR216 shows ages from 39.90 – 41.28 Ma, with comparable errors of 0.05-0.07 Ma. We averaged these ages, leading to an age of 40.60 ± 0.06 Ma. We obtained only one age for IR215, of 39.34 ± 0.54 Ma. IR119 gave two ages of 39.46 ± 0.06 and 39.71 ± 0.05 Ma, which we averaged to 39.59 ± 0.06 Ma.

Our results show that Ar-Ar dating on Eocene igneous rocks in the northwest of Iran is not straightforward, as the samples from the northwest of Iran and Azerbaijan suffered from severe excess argon, similar to previous work (van der Boon et al., 2017; Vincent et al., 2005). Replicate analyses are not always reproducible and samples with higher intensities are systematically (slightly) younger than samples yielding lower intensities. We assume that the low-intensity offset can be explained by (underestimation of) systematic uncertainties in our blanks and blank correction procedures. All obtained age plateaus, however, fall within an age range of ~36-45 Ma, i.e. around the middle Eocene.