Predicted diversity in water content of terrestrial exoplanets orbiting M dwarfs

Tadahiro Kimura and Masahiro Ikoma

Exoplanet surveys around M dwarfs have detected a growing number of exoplanets with Earth-like insolation. It is expected that some of those planets are rocky planets with the potential for temperate climates favourable to surface liquid water. However, various models predict that terrestrial planets orbiting in the classical habitable zone around M dwarfs have no water or too much water, suggesting that habitable planets around M dwarfs might be rare. Here we present the results of an updated planetary population synthesis model, which includes the effects of water enrichment in the primordial atmosphere, caused by the oxidation of atmospheric hydrogen by rocky materials from incoming planetesimals and from the magma ocean. We find that this water production in the primordial atmosphere is found to significantly impact the occurrence of terrestrial rocky aqua planets, yielding ones with diverse water content. We estimate that 5-10% of the planets with a size of $<1.3R_⊕$ orbiting early-to-mid M dwarfs have appropriate amounts of seawater for habitability. Such an occurrence rate would be high enough to detect potentially habitable planets by ongoing and near-future M-dwarf planet survey missions.

The Earth’s temperate climate has been maintained through the geochemical carbon cycle in which weathering plays a key role. On the present-day Earth, weathering takes place efficiently on land. Lands exist on the Earth because the planet has a moderate amount of seawater accounting for 0.023% of the planet’s total mass. On planets having tens of times more seawater than the Earth, weathering could not work efficiently enough to maintain temperate climates (for example, refs. 6,7). Regarding the origin of seawater, a widespread idea is that the Earth’s seawater was brought by water-laden or icy planetesimals (for example, ref. 5). Based on this idea, one would naturally predict a bimodal distribution of water content between the regions interior and exterior to the snowline in a protoplanetary disc, which was demonstrated elsewhere 6-8, as described above.

Alternatively, water can be secondarily produced in a primordial atmosphere of nebular origin through the reaction of atmospheric hydrogen with oxidizing minerals from the magma ocean, which is formed because of the atmospheric blanketing effect, thereby enriching the primordial atmosphere with water. By assuming effective water production, we recently showed that nearly Earth-mass planets can acquire sufficient amounts of water for their atmospheric vapour to survive in harsh ultraviolet (UV) environments around pre-main-sequence M stars 6. The results suggest that including this water production process significantly affects the predicted water amount distribution of exoplanets in the habitable zone (HZ) around M dwarfs.

Our planetary population synthesis model, which follows the evolution of planets’ masses, radii and orbits by combining empirical laws for several components of the planet formation process based on the planetesimal accretion hypothesis, is similar to those discussed in other work 9-11 but includes the movement of snowline location due to the thermal evolution of the protoplanetary disc, the accumulation of atmospheres of nebular origin (that is, the primordial atmospheres) also before the core mass becomes critical and the effects of water production in the primordial atmospheres. Furthermore, we have updated the treatment of some of the processes involved in planet formation/evolution according to improved understandings (Methods). In this study, we have performed Monte Carlo calculations for $1 \times 10^7$ yr with 10,000 different initial conditions for a given set of input parameters (Methods), and investigated the frequency distribution of water content among the synthesized planets.

Results

Planetary mass and semi-major axis distribution. First, our planetary population synthesis calculations demonstrate that water enrichment in the primordial atmosphere has a great effect on the atmospheric accumulation of low-mass planets such as sub-Earths and super-Earths (Fig. 1). The overall distributions of planetary masses $M_p$ and semi-major axes $a$ for an unenriched (Fig. 1a) and enriched (Fig. 1b) atmospheres are similar to each other. However, it turns out that sub-/super-Earths with enriched atmospheres consequently have more massive atmospheres in relatively cool regions (Fig. 1b, green symbols). This is because $H_2O$ enrichment leads to increasing the atmospheric mean molecular weight and the effective heat capacity through condensation and chemical reactions, making the atmosphere denser.

Water content distribution. Such an increase in low-mass planets with relatively massive $H_2O$-enriched atmospheres greatly affects the occurrence of aqua planets in the HZs around M dwarfs. Figure 2 shows the probability density distributions of the water mass fraction in planets with a mass of $0.3-3.0M_⊕$ in the present-day HZ around $0.3M_☉$ M dwarfs for the cases with enriched ($X_{H_2O} = 0.8$; Fig. 2a) and unenriched ($X_{H_2O} = 0$; Fig. 2b) atmospheres (Methods provides the definition of density probability). Such planets are hereafter called the habitable-zone nearly Earth-mass planets (HZ-NEMPs). Note that the present-day HZ is defined as the HZ around the star with an age of 5 Gyr. For comparison, the result

---

1Department of Earth and Planetary Science, Graduate School of Science, The University of Tokyo, Tokyo, Japan. 2Division of Science, National Astronomical Observatory of Japan, Tokyo, Japan. 3Department of Astronomical Science, The Graduate University for Advanced Studies (SOKENDAI), Tokyo, Japan. E-mail: t.kimura@eps.s.u-tokyo.ac.jp; ikoma.masahiro@gmail.com

---

https://doi.org/10.1038/s41550-022-01781-1
This is due to the shift in the location of the snowline with a c, results for cases with enriched atmospheres – at the age of 1 Gyr. 

Even for 3$, even for Earth-mass planets in the present-day habitable zone (HZ-NEMPs).

In the primordial atmosphere of 8$, Planet mass ($M_\text{p}$). Note that planets without any atmosphere are shown in black. The dashed box in each panel shows the region of nearly Earth-mass planets in the present-day habitable zone (HZ-NEMPs).

Obtained under the same assumptions and settings as discussed elsewhere is also shown (Fig. 2c). First, compared with the previous model (Fig. 2c), our planetary population synthesis models produce HZ-NEMPs with a much wider range of water contents even for $X_{\text{H}_2\text{O}} = 0$ (in particular, those with a water mass fraction of $\lesssim 10\%$). This is due to the shift in the location of the snowline with the protoplanetary disc’s thermal evolution, which is not included in the previous model. Initially, when viscous heating dominates, the snowline is located at $\sim 1$ au. As viscous heating diminishes and then stellar irradiation dominates, the snowline migrates inward, reaching $\sim 0.2$ au, corresponding to the outer edge of the present-day HZ on a timescale of Myr (Supplementary Fig. 1 and Supplementary Section 1.1). Here our model assumes that once the snowline passes through inward, water vapour immediately condenses onto rocky planetesimals, making planetesimals with an ice–rock ratio of unity. Thus, planets originally formed in this region can acquire small amounts of ice, depending on the timescale of planetary mass growth, orbital migration and snowline migration. The effects of this assumption are discussed in Supplementary Section 1.2 and Supplementary Fig. 2.

Water enrichment in the primordial atmosphere, which is of special interest in this study, is found to bring about a further increase in the water contents of HZ-NEMPs (Fig. 2a). Because of this enrichment, even without capturing icy planetesimals, the rocky planets obtain water from their primordial atmosphere. As demonstrated in Fig. 1b, water enrichment enhances the accumulation of primordial atmospheres. Consequently, the captured $\text{H}_2/\text{He}$ and secondarily produced $\text{H}_2\text{O}$ are sufficiently large in amount to survive the subsequent atmospheric photoevaporation process. This mechanism forms rocky planets with a water mass fraction of $< 1\%$ inside the snowline. In contrast to the previous models that predict the absence of HZ-NEMPs with Earth-like water content ($2.3 \times 10^{-4}$), our model with enriched primordial atmospheres predicts that a significant number of such HZ-NEMPs are formed. It is noted that the abundance of HZ-NEMPs with Earth-like water contents is significantly affected by the protoplanetary disc conditions, especially the disc lifetime. Most of the HZ-NEMPs with small amounts ($\lesssim 1$ wt%) of water are found to form in discs with a lifetime of $\lesssim 3$ Myr. In longer-lived discs, the outer icy planets significantly migrate inward, pushing the inner rocky planets closer to the disc inner edge. As a result, only ice-dominant planets exist in the HZ.

Dependence on central stellar mass. Similarly to Fig. 2, Fig. 3 shows the probability density distribution for HZ-NEMPs orbiting stars of three different masses ($0.1$, $0.3$ and $0.5M_\odot$), around which the HZs are located at $0.03–0.07$, $0.10–0.20$ and $0.20–0.40$ au, respectively, in their main-sequence phase. Evidently, the larger the stellar mass, the higher the abundance of HZ-NEMPs with small water content ($\lesssim 1$ wt%). This is due to two effects: first, the region that corresponds to the HZ when the host star is on its main sequence is located at...
a larger orbital distance for larger stellar mass. In that region of a protoplanetary disc, therefore, protoplanets can grow larger around the more massive stars; consequently, the protoplanets can obtain more atmospheric gas, thereby producing more water. Second, the less massive the host star, the higher the stellar extreme-ultraviolet (XUV) irradiation in the HZ during the pre-main-sequence phase. Therefore, the planets in the HZ undergo severer loss of atmosphere and are less probable to keep water. In cases with unenriched atmospheres, the distribution also differs depending on the stellar mass (dashed bars). The distributions for 0.1 and 0.3\(M_\odot\) stars are similar to each other, whereas the planets around 0.5\(M_\odot\) stars have a much wider range of water contents, and the distribution is rather similar to that for enriched atmospheres. This indicates that many HZ-NEMPs around 0.5\(M_\odot\) stars have obtained relatively large amounts of icy planetesimals during their formation. This is because the snowline comes closer to the HZ around more massive stars; for 0.5\(M_\odot\) stars, the former comes inside the latter.

Discussion

Our planetary population synthesis calculations have shown that HZ-NEMPs orbiting M dwarfs can have diverse water contents by the effects of the disc’s thermal evolution and water enrichment in the primordial atmosphere. In particular, the latter effect enables about 1\% HZ-NEMPs to have water amounts comparable to Earth oceans, provided the atmospheric water mass fraction of 80\% is kept. In particular, the latter effect enables about 1\% HZ-NEMPs to have water amounts comparable to Earth oceans, provided the atmospheric water mass fraction of 80\% is kept. In particular, the latter effect enables about 1\% HZ-NEMPs to have water amounts comparable to Earth oceans, provided the atmospheric water mass fraction of 80\% is kept.

The ratio \(P_{\text{HZ}}/P_{\text{HZ,0}} = 1\) corresponds to about 80wt\% H\(_2\)O in the H\(_2\)-He-H\(_2\)O atmosphere with solar He/H ratio of 0.385 (ref. \(^{10}\)). Thus, the assumed value of \(X_{\text{HZ}} = 0.8\) is feasible if the entire atmospheric hydrogen is equilibrated with oxides. We should note that our results and conclusion hardly change for \(X_{\text{HZ}} \geq 0.5\), because the atmospheric mass significantly increases in that range\(^{1}\). We further discuss the implication of this water production process in the primordial atmosphere for the Earth in Supplementary Section 2.3.

Our planetary population synthesis model is based on the planetesimal accretion scenario, and does not consider pebbles. Pebble accretion may affect our results and conclusion mainly in the following three ways: changes in the disc gas composition, changes in the oxidation state of the planetary rocky components and changes in the planetary mass distribution. First, drifting pebbles can enhance the water content of disc gas inside the snowline through the sublimation of ice\(^{41}\). However, this change in the disc gas composition would hardly affect the total mass of the water produced in the primordial atmosphere when the atmosphere is equilibrated with the magma ocean. Second, pebbles drifting from outside the snowline would be more oxidized than solids originally existing inside the snowline\(^{42}\). Since the equilibrium partial pressure of the water vapour in the atmosphere is higher for more oxidized magma, the accretion of oxidized rocky components would work in favour of the water production and formation of planets with small water content. Finally, a pebble accretion scenario may significantly change the resultant \(M_\text{–}\)a distribution of the synthesized planets\(^{42}\). Since the water amount in the atmosphere largely depends on the planetary mass, the mass distribution of the planets in the HZ affects the water amount distribution. Further studies including pebbles are needed to quantify these effects.

Our models place constraints on the radii of temperate aqua planets, which will be useful for ongoing and future exploration of habitable planets around M dwarfs. Figure 4 shows the relationships between the radius and water content of HZ-NEMPs with enriched atmospheres orbiting 0.3\(M_\odot\) M dwarfs. Most of the HZ-NEMPs of \(>1.3R_\oplus\) are shown in bluish colours, indicating that those planets have thick hydrogen-rich atmospheres (\(\geq 10^3M_\oplus\)). Given the fact that thick atmospheres bring about such a strong blanketing effect that an H\(_2\)O layer below the atmosphere would be, if any, in a super-critical state, those planets are unlikely to be habitable. The other HZ-NEMPs of \(>1.3R_\oplus\) (black symbols) retain large amounts of water (\(\geq 0.1M_\odot\)) and are far from Earth like (but worth atmospheric characterization with JWST\(^{17}\) and Ariel\(^{18}\) for verifying our theoretical prediction).

The HZ-NEMPs of 0.7–1.3\(R_\oplus\) (Fig. 4, black) have completely lost their hydrogen atmospheres, ending up with rocky planets covered with oceans. It turns out that those planets are diverse in water content and do include planets with Earth-like water content. Several climate studies argue the amounts of seawater that are appropriate for temperate climates, considering the effects of seafloor weathering, high-pressure ice, water cycling and heterogeneous surface water distribution\(^{14,19–21}\). According to those studies, the appropriate seawater amount ranges from ~0.1 to 100.0 times that of the Earth. From an observational point of view, it would be important to exclude planets unlikely to be habitable in advance. Among those HZ-NEMPs, there are relatively low-mass rocky planets that have deep oceans with high-pressure ice; such planets are unlikely to have temperate climates\(^{14,25}\). Meanwhile, one can identify rocky planets with ocean mass fractions \(\geq 100\) times the present-day Earth’s one, which are also probably uninhabitable, if the planetary masses and radii are measured within \(\leq 20\%\) and 5\% accuracies, respectively. In Fig. 4, about 25\% HZ-NEMPs of 0.7–1.3\(R_\oplus\) are such ocean-dominated planets.

The remaining 75\% would be identified as ‘water-poor rocky planets’ (that is, rocky planets without hydrogen-rich atmospheres or thick oceans), which are capable of having Earth-like temperate climates as long as they have little amounts of seawater (\(\leq 20.001\%)\(^{19,21,27}\).
Thus, excluding the completely dry planets, which account for about 95% of those remaining HZ-NEMPs, the HZ-NEMPs with appropriate amounts of seawater for habitability are estimated to account for ~5% of the ‘water-poor rocky planets’ orbiting 0.3–0.6 M\_\_ M\_ stars, for example, more than 10% of the water-poor rocky planets are expected to have the appropriate amounts of seawater. Note that the appropriate water amount for temperate climate can be even wider, considering the possibility of water sequestration in the mantle and of further water loss due to the absence of silicate weathering. On the other hand, for tidally locked planets, water can be trapped as ice on the night side, leading to atmospheric cooling, 

\[ \frac{M_{\text{water}}}{M_{\text{terrestrial}}} = 10^{-3} \]  

for drawing purposes.

**Methods**

Here we describe the details about our planetary population synthesis model, which is based on the planetesimal-driven core accretion scenario. The model includes the evolution of the central star; the evolution and radial structure of the protoplanetary disk, as well as the growth of solid cores; the accumulation and loss of atmospheres; orbital migration; and dynamical interactions between protoplanets, including resonance trapping and orbital instabilities for multi-protoplanet systems and their outcomes. In addition, we describe the calculation method for the radius of the isolated planet during its thermal evolution and the initial conditions and parameters for generating the planetary populations. We discuss the effects of the updated components of the model from a previous study and the consistency with the observed exoplanet populations in Supplementary Sections 2.1 and 2.2 and Supplementary Figs. 3 and 4.

**Stellar evolution.** From the table provided in another work, we take the stellar radius, luminosity and effective temperature, as a function of stellar mass \( M \) and age for the solar metallicity with He abundance of \( Y = 0.275 \). These data are used for calculating the protoplanetary disc temperature, planetary equilibrium temperature and atmospheric photoevaporation rate in our planetary population synthesis calculations.

**Disc gas profile and evolution.** (1) Initial profiles

The initial profile of the disc gas surface density is assumed as 

\[ \Sigma(I) = \Sigma_0 \left( \frac{r}{r_0} \right)^{-q} \exp \left[ -\left( \frac{r}{r_0} \right)^{-q/2} \right] \times \left( 1 - \frac{q}{2} \right) \]

where \( \Sigma_0 \) is the gas surface density at orbital radius \( r = r_0 \), \( r_0 \) is the disc gas characteristic radius beyond which the surface density decays exponentially and \( r_c \) is the inner edge radius. This is the so-called self-similar solution with turbulent viscosity \( \nu_{\text{turb}} \propto r^{3/2} \) (refs. [25, 26]), including the smooth cutoff at \( r = r_c \). We set \( q = 0.9 \), which is inferred from the observations of protoplanetary discs, and \( r_c = 1 \) au. The surface density at \( r = r_0 \) namely, \( \Sigma_0 \), is calculated from the total disc gas mass \( M_{\text{disc}} \) as

\[ \Sigma_0 = \frac{2}{\pi} M_{\text{disc}} \frac{r_c^{-q}}{r_c^{q-2}} \frac{1}{r_c^2} \]

This is obtained by integrating equation (1) multiplied by \( 2\pi r \) (that is, \( 2\pi r \Sigma(I) = 0 \)) from \( r = 0 \) to \( r_c \), ignoring the exponential and inner-edge cutoff terms.

(2) Viscous diffusion

The evolution of the disc gas surface density \( \Sigma \) is assumed to occur via radial viscous diffusion, photoevaporation and absorption by embedded planets and is expressed as 

\[ \frac{\partial \Sigma}{\partial t} = -\frac{3}{2} \frac{\partial}{\partial r} \left( r^{1/2} \Sigma \frac{\partial \Sigma}{\partial r} \right) \]

where \( -\Sigma_{\text{pe}} \) and \( -\Sigma_{\text{planet}} \) are the sink terms due to photoevaporation and absorption by planets, respectively. We adopt the \( \alpha \)-prescription for turbulent viscosity:

\[ \nu_{\text{turb}} = \alpha c_s \Sigma \]

where \( c_s = \sqrt{k_B T_{\text{gas}}/(\mu m_\text{H})} \) is the isothermal sound speed and \( T_{\text{gas}} = T_\text{g,II}/c_s^2 \) is the disc scale height. Here \( \mu \) is the mean molecular weight, which is assumed to be 2.34 for gas with solar abundances; \( T_{\text{g,II}} \) is the midplane temperature (calculated below); \( k_B \) and \( m_\text{H} \) are the Boltzmann constant \((1.38 \times 10^{-24} \text{erg K}^{-1})\) and hydrogen atomic mass \((1.66 \times 10^{-24} \text{g})\), respectively; and \( k_B \) is the Keplerian frequency. We treat \( \alpha_{\text{visc}} \) as an input parameter.

\[ \Sigma = \Sigma_0 (r_0/r)^{-q} \exp \left[ -\left( r/r_0 \right)^{-q/2} \right] \]

Equation (3) is solved with a log-uniform grid with 500 points \( (N_{\text{grid,disc}}) \) extending from \( r_0 \) to \( r_{\text{in,j}} = 1 \) au. The boundary conditions are \( \Sigma(r_{\text{in,j}}) = 0 \) and \( \Sigma(r_{\text{in,j}}) = 0 \) for \( r < r_{\text{in,j}} \) and \( r > r_{\text{in,j}} \), respectively.

(3) Photoevaporation

We consider photoevaporation processes caused by UV irradiation from the central star (internal photoevaporation) and from nearby massive stars (external photoevaporation). The total photoevaporation rate \( \dot{M}_{\text{pe}} \) is determined by the sum of the two photoevaporation rates. The external photoevaporation rate is calculated as per another work. The model assumes that the far-ultraviolet photons from nearby massive stars uniformly evaporate the disc gas only in the regions exterior to the effective gravitational radius. The gravitational radius for the dissociated gas is defined as

\[ r_\text{g} = \frac{GM_{\text{disc}}}{c_s^2} \]

where \( G \) is the gravitational constant \((6.67 \times 10^{-11} \text{cm}^3 \text{g}^{-1} \text{s}^{-2})\); \( c_s \) is the speed of sound at a temperature of \( 1 \times 10^4 \) K and mean molecular weight of 1.35 \((\text{ref. [27]}\)), given that the hydrogen is fully dissociated and the He/H ratio is equal to the solar value. Assuming that the disc gas is uniformly removed only from the regions exterior to the effective gravitational radius \( r_\text{g} \), the external photoevaporation rate is given by 

\[ \dot{M}_{\text{pe,ext}} = \begin{cases} 0 & \text{for } r < r_\text{g} \\ -\frac{M_{\text{disc}}}{r_{\text{in,j}}^2} & \text{for } r > r_\text{g} \\ \end{cases} \]

where we assume \( r_{\text{in,j}} = 1 \) au and regard \( M_{\text{disc}} \) as an input parameter providing the total mass-loss rate. For the effective gravitational radius, analytical estimates and numerical results show that \( \beta = 0.1–0.2 \) would be appropriate. Here we set \( \beta = 0.14 \) \((\text{ref. [28]}\).

Next, the internal photoevaporation rate is calculated from the following equation:

\[ \dot{M}_{\text{pe,wind}}(r) = \begin{cases} 0 & \text{for } r < \beta r_\text{g} \\ \frac{2\pi c_s^2 \rho_{\text{wind}}(r)}{M_{\text{disc}}} & \text{for } r > \beta r_\text{g} \end{cases} \]

Here \( \beta r_\text{g} \) and \( c_s \) are the gravitational radius and the speed of sound for ionized gas at a temperature of \( 1 \times 10^4 \) K and mean molecular weight of 0.68 (assuming fully

**Protoplanetary disc.** Our planetary population synthesis calculations need the surface density distributions of gas and solids (or planetesimals) in the protoplanetary disc and their evolution, as well as need the disc midplane temperature. These are calculated in a similar way to that done in another work, which is summarized below.
where $n_i$ is the enhancement factor associated with H$_2$O condensation (as described below) and $\Sigma_{\text{rock}}^{i(0)}$ is the initial surface density of rocks, which is given by

$$\Sigma_{\text{rock}}^{i(0)} = \eta_{\text{rock}} \Sigma_{\text{rock}}^{\text{inc}}.$$

Here $\Sigma_{\text{rock}}$ is the initial surface density of rocks at $r=r_0$, and $r_{\text{rock}}$ is the characteristic radius of the existence region of planetesimals (the so-called planetesimal disc). We set $q_i=1.5$ (refs. 46, 47) and $r_{\text{low}}=0.5a_{\text{in}}$, (ref. 4). Similar to $\Sigma_{\text{g}}$, $\Sigma_{\text{rock}}^i$ is calculated by

$$\Sigma_{\text{rock}}^i = \frac{2}{\pi} \eta_{\text{rock}} M_{\text{init}} a_{\text{rock}}^{-1} r_{\text{rock}}^{-2}.$$

The temperature due to indirect irradiation $T_{\text{irr}}$ is given by

$$T_{\text{irr}}^4 = \frac{1}{\pi} \frac{3}{8} R^2 + \frac{1}{2\pi} \frac{\tau}{r} \tilde{E},$$

$\tilde{E}$ is the viscous energy generation rate per unit area surface. The Rosseland mean opacity $\kappa_\text{R}$ is given by the analytical fitting of the opacity of dust grains provided in another work44. We ignore the gas opacity and simply set $\kappa_\text{R}=0$ cm$^2$ g$^{-1}$ after the evaporation of dust grains. Also, we simply set $\tau_\text{f}=\max(2.4R^2,0.5)$ (refs. 44, 45). The minimum value of 0.5 is adopted so that the coefficient of $E$ in equation (11) approaches unity in the optically thin limit. The viscous heating rate $\dot{E}$ is

$$\dot{E} = \frac{3}{2} \frac{H_{\text{visc}}} {\tau_{\text{in}}} \tilde{E}.$$
the planet and damping due to the disc gas drag. Following the analytic estimates provided elsewhere, it is expressed as

$$\dot{\rho}_{\text{plan}} = 4.1 \frac{(\rho_{\text{gas}})}{10^{-7} \text{g cm}^{-3}} \frac{1}{10^{-7} \text{g cm}^{-3}} \frac{1}{2/15} \frac{(\rho_{\text{gas}})}{10^{-15} \text{g cm}^{-3}} \frac{2}{15} \frac{1}{2/15},$$

(27)

where $\rho_{\text{gas}} = 3.0 \text{g cm}^{-3}$ and $m_{\text{gas}} = 10^9 \text{g}$ are the material density and mass of planetesimals. Finally, we calculate the effective capture radius enhanced by the atmospheric gas drag $R_{\text{cap}}$ (ref. 57) as

$$R_{\text{cap}} = \frac{3 \sqrt{\rho_{\text{gas}}}}{2} \frac{2 \sqrt{2} \rho_{\text{gas}}}{2GM_{\text{cap}}/R_{\text{cap}}} \rho_{\text{gas}},$$

(28)

where $R_{\text{cap}}$ is the planetesimal's radius derived from $m_{\text{gas}}$ and $\rho_{\text{gas}}$. $\nu_{\text{esc}} = \sqrt{\frac{2GM_{\text{cap}}}{R_{\text{cap}}}}$ is the random velocity of the planetesimals and $\rho_{\text{gas}}$ is the envelope gas density at radius $R_{\text{cap}}$. After disc gas dissipation, some of the planetesimals that encountered a planet are ejected from the system, instead of colliding with the planet. This ejection rate can be estimated by comparing the collisional cross section and the scattering cross section, and is given by

$$M_{\text{eject}} = \left( \frac{\nu_{\text{esc}}}{\nu_{\text{gas}}} \right)^2 M_{\text{core}},$$

(29)

where $\nu_{\text{gas}} = \sqrt{\frac{2GM_{\text{cap}}}{R_{\text{cap}}}}$ is the escape velocity from the host star and $\nu_{\text{esc}} = \sqrt{\frac{2GM_{\text{cap}}}{R_{\text{esc}}}}$ is the surface velocity of the planet. This is used for the calculation of evolution of $L_{\text{rad}}$ (equation (21)).

### Atmospheric accumulation and loss. Purely hydrostatic equilibrium phase. During vigorous planetesimal accretion, the atmospheric mass is rather small because of large energy deposition. Then, the atmosphere is in hydrostatic equilibrium and thermal steady state. We solve the standard set of equations for stellar structure, namely,

$$\frac{\partial P}{\partial P} = \frac{GM_{\text{core}}}{4\pi R^2},$$

(30)

$$\frac{\partial T}{\partial M_{\text{gas}}} = -\frac{GM_{\text{gas}}}{4\pi R^2} \nabla,$$

(31)

$$\frac{\partial R}{\partial M_{\text{gas}}} = \frac{1}{4\pi R^2},$$

(32)

where $P, T$ and $\rho$ are the pressure, temperature and density of the atmosphere (or envelope) gas, respectively; $R$ is the radial distance to the planet's centre; and $M_{\text{gas}}$ is the total mass inside the sphere of radius $R$. In addition to the above equations, we use the ideal state of equation (or the $P-T$ relationship) for chemical equilibrium mixtures composed of $H$ and $O$-bearing molecules and $He$, taking $H_2O$ sublimation/condensation into account. Also, $V$ is the temperature gradient, $d\log T/d\log P$, for radiative diffusion or convection (dry or moist adiabat). In the early stages of planetary accretion, the temperature is high enough at the bottom of the atmosphere to maintain a global magma ocean (ref. 8,9). The atmosphere–disc and $\tau_{\text{cool}}$ are related in the form of $dT$, which is assumed to be

$$\tau_{\text{cool}} = \tau_{\text{damp}} = \tau_{\text{esc}} = \frac{L_{\text{rad}}}{L_{\text{esc}}},$$

(33)

where $L_{\text{esc}}$ and $L_{\text{rad}}$ are the luminosity values due to planetesimal accretion, solid core cooling and radioactive decay, respectively. The accretion luminosity is

$$L_{\text{ac}} = \frac{GM_{\text{core}} C_{\text{rock}} R_{\text{esc}}}{R_{\text{core}},}$$

(34)

where $M_{\text{core}}$ is the planetesimal accretion rate and $R_{\text{esc}}$ is the solid planet radius. By assuming that the core surface temperature $T_{\text{surf}}$ and disc gas temperature $T_{\text{esc}}$ are related in the form of $T_{\text{surf}} = T_{\text{esc}}/4$, as indicated by the analytical solution of the fully radiative atmosphere, $L_{\text{cool}}$ is given by

$$L_{\text{cool}} = \frac{M_{\text{core}} C_{\text{rock}} T_{\text{esc}}}{4 \text{ disc}},$$

(35)

where $C_{\text{rock}}$ is the specific heat of rock (1.2 x 10$^6$ erg g$^{-1}$ K$^{-1}$) and $t_{\text{esc}}$ is the disc dissipation timescale. We set $t_{\text{esc}} = 1 \times 10^7$ yr in our simulations, since disc dissipation occurs on a timescale of $10^4$--$10^7$ yr due to photoevaporation in the final stage of disc evolution. Increasing this value by an order of magnitude has little effect on our results. Although the above equation was derived assuming a thin radiative atmosphere, the vapour-rich primordial atmosphere considered in this study is thick even in the final stage of disc evolution, and the lower layer is often convective. Even in that case, the changing rate of temperature at the bottom of the atmosphere is about the same as that at the radiative–convective boundary. Therefore, equation (35) is still considered to be a good approximation for describing the cooling luminosity. The radiogenic luminosity $L_{\text{rad}}$ is simply set to $2 \times 10^{10} (M_{\text{core}}/M_{\text{planet}}) \text{ erg s}^{-1}$ (ref. 10).

### Quasi-static contraction phase. Once planetesimal accretion stops, the vapour-mixed atmosphere gravitationally contracts and further accumulation of disc gas occurs. Thus, we have to integrate the equation of entropy change, in addition to equations (30)–(32) as

$$\frac{dL_{\text{gas}}}{dM_{\text{gas}}} = e - T \frac{dS}{dT},$$

(36)

where $L_{\text{gas}}$ is the total energy flux passing through a spherical surface of radius $R$ (or luminosity), $e$ is the energy generation rate per unit mass and $S$ is the specific entropy. We assume that the accumulating disc gas composed predominantly of $H$ and $He$ never mixes with the lower vapour-mixed layer and thus the mass of the vapour-mixed atmosphere ($M_{\text{mix}}$) is conserved in this phase. Hereafter, the accumulated mass of disc gas is denoted by $M_{\text{total}}$. Therefore, we calculate the quasi-static contraction of the two-layer atmosphere in this phase.

The numerical integration of equation (36) is done based on the total energy conservation approximation adopted in several previous studies on the formation of gas giants (ref. 14–16). For the total atmospheric energy conservation between time $t - \Delta t$ and $t$ being considered, the following relation holds (ref. 17) provides the derivation:

$$\frac{E_{\text{con}}(t) - E_{\text{con}}(t - \Delta t)}{\Delta t} = L_{\text{con}} + e_{\text{gas}} \left[ M_{\text{con}}(t) - M_{\text{con}}(t - \Delta t) \right] - L,$$

(37)

or

$$\Delta t = \frac{E_{\text{con}}(t) - E_{\text{con}}(t - \Delta t) - e_{\text{gas}} \left[ M_{\text{con}}(t) - M_{\text{con}}(t - \Delta t) \right]}{L_{\text{con}} - L}.$$

(38)

where $e_{\text{gas}}$ is the total energy per unit mass (that is, the sum of specific internal energy and gravitational energy) of the disc gas at the outer boundary,

$$M_{\text{total}} = M_{\text{core}} + M_{\text{mix}}$$

is the total envelope mass and $E_{\text{con}}$ is the envelope's total energy defined by

$$E_{\text{con}} = \int_{m_{\text{total}}}^{M_{\text{total}}} \left( u - \frac{GM_{\text{core}}}{r} \right) dm.$$
where coefficient $D$ is empirically given by
\[ D = 0.29 \left( \frac{M_{\text{disc}}}{M_p} \right)^{-2} \left( \frac{M_p}{M_\star} \right)^{4/3} \Omega_\star^2 \delta a, \]
and $\Sigma_{\text{gpp}}$ is the gas surface density at the bottom of the gap, which is also empirically given by
\[ \Sigma_{\text{gpp}} = \frac{\Sigma_p}{1 + 0.04 K}. \]
where $K$ is given by
\[ K = \left( \frac{M_p}{M_\star} \right)^3 \left( \frac{h}{a} \right)^5 \alpha_n^{0.1}. \]

Here $\alpha_n$ is a parameter for disc turbulent viscosity, which is not necessarily the same as $\alpha_{\text{vis}}$. When wind-driven accretion is dominant in the global angular momentum transfer, $\alpha_n$ is larger than $\alpha_{\text{vis}}$ by about one order of magnitude. Thus, we set $\alpha_n = 0.1 \alpha_{\text{vis}}$ in this study.

Equation (41) is valid if there is sufficient disc gas outside the gap. As the disc depletes, the gas accretion rate is limited by the total disc mass accretion rate $M_{\text{disc}}$ given by
\[ M_{\text{disc}} = 6\pi H^2 \frac{\partial}{\partial r} \left( \Sigma_{\text{gpp}} F_{\text{acc}} \right). \]

Noting that solid accretion never occurs in this phase.

**Atmospheric thermal evolution and loss.**

**Atmospheric thermal evolution.** The planetary radius $R_\star$ is expressed by
\[ R_\star = R_{\text{core}} + \Delta R_{\text{env}}, \]
where $R_{\text{core}}$ is the solid core radius and $\Delta R_{\text{env}}$ is the thickness of the envelope.

The solid core is assumed to consist of iron, silicate and, if present, ice. Following another study, we calculate the core radius as
\[ R_{\text{core}} = R_{\text{rock}} (1 + 0.55 f_{\text{ice}} - 0.14 f_{\text{Si}},) \]

where $f_{\text{ice}}$ is the ice mass fraction and $R_{\text{rock}}$ is the pure rocky (iron + silicate) core radius, which is calculated as
\[ R_{\text{rock}} = \left( 0.0592 f_{\text{rock}} + 0.0975 \right) (\log [M_{\text{core}}] / 2) + (0.2337 f_{\text{rock}} + 0.4938) \log [M_{\text{core}}] + (0.3102 f_{\text{rock}} + 0.7932), \]

where $f_{\text{rock}}$ is the Si/(Si + Fe) mass ratio set to the Earth-like value of 0.66 in this study.

To evaluate the planetary radius including the envelope, we calculate the quasi-static thermal evolution of the planet after disc gas dissipation. For numerical convenience, we treat the upper part of the envelope (simply called the atmosphere) as a quasi-static evolution of the planet after disc gas dissipation. For numerical convenience, we treat the upper part of the envelope (simply called the atmosphere) as a quasi-static evolution of the planet after disc gas dissipation.

The outer boundary condition for the envelope structure is given in terms of temperature and pressure, respectively, by
\[ P = P_{\text{out}}, \quad T = T_{\text{out}} \] at $M_{\text{env}} = M_p$.

To evaluate $R_{\text{env}}$ and $T_{\text{out}}$, we calculate the radiative–convective structure of the atmosphere in the same way as that done elsewhere. The atmosphere is assumed to be plane parallel and its mass and thickness are negligible compared with the total planetary mass and radius, respectively. Thus, the gravitational acceleration $g = G M_\star / R_{\text{env}}^2$ is constant throughout the atmosphere, where $R_{\text{env}}$ denotes the radius at the boundary between the atmosphere and envelope (note that it is not equal to the planetary radius $R_\star$, as shown below). This boundary is assumed to be located at the radius where the optical depth for the stellar visible radiation, $\tau_{\text{vis}}$, is 10 such that the temperature gradient in the envelope is hardly affected by the irradiation from the central star.

The temperature in the radiative region is calculated by the following analytical formula:
\[ \sigma T^4 = \frac{F_p}{2} + \frac{\alpha_\sigma}{2} \left[ 1 + \frac{3}{2} \gamma (\frac{\tau_{\text{vis}}}{\tau_{\text{opt}}}) \right] \] for $\tau_{\text{vis}} > \tau_{\text{opt}}$.

The radiative–convective boundary (that is, the tropopause) is determined so that the temperature and radiation flux continuously connect, using the same numerical procedure described elsewhere. As a result, $R_{\text{env}}$ and $T_{\text{out}}$ are determined when $L_\star$ and $R_{\text{env}}$ are given from the envelope structure calculation. Thus, we find a self-consistent set of values ($P_{\text{out}}$, $T_{\text{out}}$, $L_{\text{env}}$, $L_\star$, $R_{\text{env}}$) at time $t_0$ for a given $M_p$ with iterations. Then, we define the planetary radius $R_\star$ as the pressure level of 10 mbar, and the envelope thickness is derived simply by $\Delta R_{\text{env}} = R_\star - R_{\text{env}}$.

If the planet has experienced the runaway accretion phase, however, the thick envelope with nebular composition exists instead of the water-enriched envelope. The radius in this case is calculated with the following fitting formula:
\[ R_{\text{env}} = \frac{2.06 \mu_{\text{env}} / \mu_{\text{rock}}}{\delta a} (\frac{R_\star}{T_\star})^{0.21} \times (\frac{\mu_{\text{env}}}{\mu_{\text{rock}}})^{0.99} \] for $\frac{R_\star}{T_\star} < 0.1$ and $\Delta R_{\text{env}} = \frac{9 \kappa_1 T_{\text{env}}}{\mu_{\text{env}} \lambda_1}$.

where $g$ is the gravitational acceleration and $\mu_{\text{env}} = 2.34$ is the mean molecular weight of nebular gas.

**Atmospheric photoevaporation.** After disc dispersal, atmospheric escape (or photoevaporation) occurs because of stellar XUV irradiation. We assume that the atmospheric photoevaporation starts once the radial optical depth at the planet's location measured from the central star, $\tau_\star$, becomes less than unity. We use the fitting formula given elsewhere for the escape rate, which yields higher values for the escape rate than those for the energy-limited escape rate when the escape parameter is small ($\ll 10$). Although the atmospheres in our calculations are enriched with water vapour, we simply adopt the formula derived for hydrogen-dominated atmospheres. Since the enrichment with water leads to raising the value of the escape parameter and thus reducing the escape rate, our calculations are expected to overestimate the escape rate. The fitting formula is expressed in the form of
\[ M_{\text{env}} = \frac{\dot{F}_{\text{XUV}}}{\sqrt{L_{\text{XUV}}}} \left( \frac{\alpha_0}{\mu_0} \right)^{0.1} \left( \frac{\alpha_0}{\mu_0} \right)^{0.1} \left( \frac{\alpha_0}{\mu_0} \right)^{0.1} \]

where $\beta$, $\alpha_0$, $\alpha_1$, and $k$ and $\kappa$ are fitting coefficients (values provided in another work). $\dot{M}_{\text{env}}$ is the incoming XUV flux and $\dot{A}$ is the escape parameter. The stellar XUV luminosity, $L_{\text{XUV}}$, is taken from a table given in another study. Since that table only gives the best-fit relation between stellar $L_{\text{XUV}}$ and time, we set
\[ L_{\text{XUV}} = 10^7 \text{erg cm}^{-2} \text{ s}^{-1} L_{\text{XUV}}^{\text{J}} \] to account for the variation between the observed stars. Here $L_{\text{XUV}}^{\text{J}}$ is the best-fit value of XUV luminosity given in the table, and $\delta$ is the deviation factor that follows the normal distribution with $\mu = 0$ and $\sigma = 0.359$.

**Orbital migration.** The planetary orbital migration rate can be calculated by
\[ \frac{\Delta a}{dt} = \frac{2 \gamma}{M_p \mu_{\text{rock}} \Omega_\star} \]

where $\gamma$ is a parameter for disc turbulent viscosity, which is not necessarily the same as $\alpha_{\text{vis}}$. When wind-driven accretion is dominant in the global angular momentum transfer, $\alpha_n$ is larger than $\alpha_{\text{vis}}$ by about one order of magnitude. Thus, we set $\alpha_n = 0.1 \alpha_{\text{vis}}$ in this study.
where $\Gamma$ is the total torque acting on the planet. In the type-I regime, the torque is given as the sum of the Lindblad ($\Gamma_l$), co-rotation ($\Gamma_c$) and thermal ($\Gamma_t$) components, that is,

$$\Gamma = \Gamma_l + \Gamma_c + \Gamma_t.$$  

We adopt the formula for the Lindblad torque derived elsewhere\(^{76}\) through three-dimensional hydrodynamic simulations as

$$\Gamma_l = (-2.34 + 1.50 \beta_4 - 0.10 \beta_5) f(\Gamma_p) \Gamma_{\nu},$$

where $\beta_4 = \log [\Sigma_{\text{disc}}]/\log [\Sigma]$, $\beta_5 = \log [\Sigma]/\log [\rho]$, $\gamma_p$ is the thermal diffusion coefficient at the disc midplane and $\Gamma_{\nu}$ is the characteristic torque expressed as

$$\Gamma_{\nu} = \left( \frac{M_p}{M_*} \right)^2 \left( \frac{H_{\text{disc}}}{d} \right)^2 \Sigma_{\text{d}} \alpha_{\text{d}}^2 \Omega_p^2.$$  

The function $f(\Gamma_p)$ transitions from $1/\Gamma_p$ (adiabatic regime) to unity (locally isothermal regime) as $\Gamma_p$ increases (that is, the thermal diffusion of the disc gas becomes effective). Here $\gamma$ is the specific heat ratio of the disc gas. The explicit form of $\Gamma_{\nu}$ and $\alpha_{\text{d}}$ is provided elsewhere\(^{87}\).

We also use the formula for the co-rotation torque derived in another work\(^{51}\), which is expressed as the sum of four components:

$$\Gamma_c = \Gamma_{\text{Corr}} + \Gamma_{\text{Cas}} + \Gamma_{\text{Comp}} + \Gamma_{\text{Cov}}.$$  

where $\Gamma_{\text{Corr}}$, $\Gamma_{\text{Cas}}$, and $\Gamma_{\text{Comp}}$ are the torques arising from the radial gradient of vortensity, entropy and temperature, respectively. The last term, namely, $\Gamma_{\text{Cov}}$, is associated with the viscously created vortensity arising during the horseshoe $U$-turns (more detailed explanations provided elsewhere\(^{81-83}\)). Each of the first three torques is calculated by the combination of two terms, for example,

$$\Gamma_{\text{Corr}} = (1 - \epsilon_c) \Gamma_{\text{Corr,reg}} + \epsilon_c \Gamma_{\text{Corr,ateq}},$$

where $\Gamma_{\text{Corr,reg}}$ and $\Gamma_{\text{Corr,ateq}}$ are the vortensity-related components of the linear co-rotation torque and horseshoe torque, respectively. The horseshoe torque $\Gamma_{\text{Horseshoe}}$ is the product of the unsaturated horsehose drag $\Gamma_{\text{Horseshoe}}$ and saturation function $f_s$ (that is, $\Gamma_{\text{Corr,ateq}} f_s$): the latter expresses the saturation of horseshoe drag and depends on the horseshoe width. The blending coefficient $\epsilon_c$ is derived in another work\(^{51}\) by fitting their numerical results. Note that the saturation function and blending coefficient are different among $\Gamma_{\text{Corr,reg}}$, $\Gamma_{\text{Cas}}$, and $\Gamma_{\text{Comp}}$. Also, $\Gamma_{\text{Comp}}$ has no component of horseshoe torque (that is, no linear co-rotation torque) because this torque only arises from the horseshoe region. Again, the explicit forms of these torques are derived elsewhere\(^{51}\).

The thermal torque $\Gamma_t$ arises when the horseshoe gas cools during the $U$-turn by thermal radiation (cold torque) and when the horseshoe gas is heated by the luminous planet (heating torque). We calculate the thermal torque using the following analytical formulation\(^{51}\):

$$\Gamma_t = 1.61 \frac{L}{\rho} \left( \frac{H_{\text{disc}}}{L} \right) \left( \frac{L}{L_{\text{c}}} - 1 \right) \Gamma_{\nu},$$

where $q = -\beta_3/\beta_4 + 1/\gamma$, $\lambda_c = \sqrt{\gamma_p} (\sqrt{M_\ast} / \Omega_p)$, $L_p$ is the planetary luminosity and $L_{\text{c}} = 4\pi M_p \sqrt{G M_* / r}$ is the critical planetary luminosity above which the positive heating torque exceeds the negative cold torque. Note that, as discussed in another study\(^{51}\), the thermal torque effectively acts only when $M_{\text{eq}} > \gamma_p G M_p / r$ is satisfied. This condition means that the thermal diffusion timescale across the Bondi radius is shorter than the acoustic timescale. If $M_{\text{eq}} > \gamma_p G M_p / r$, we set $\Gamma_t = 0$.

When the planet grows massive enough to open a gap in the disc, its orbital migration transitions to the type-II regime. Recently, it was found\(^{51}\) that the migration rate in this regime can be expressed in a similar way as that in the type-I regime by replacing the gas surface density $\Sigma_{\text{d}}$ with $\Sigma_{\text{eq}}$ and removing the co-rotation torque. To smoothly connect both regimes, we express the total torque in the type-II regime as

$$\Gamma = \Gamma_l + (\Gamma_l + \Gamma_t) \exp(-K_{\text{K}}) + 0.94K_{\text{K}},$$

where $K = 20$ is the typical value of $K$ where the disc gas gap becomes deep enough\(^{52}\). This is similar to the formula given elsewhere\(^{51}\), but the formulae for $\Gamma_l$ and $\Gamma_t$ are different, and that for $\Gamma_c$ is also included. Since equation (61) becomes equal to equation (58) when $K$ is small (that is, the planet mass is small), we always use equation (64) in the calculation of this study.

Resonance trapping. Planets or planetary embryos with converging orbits can be captured in mean-motion resonances with each other; this phenomenon is called resonance trapping\(^{52}\). In this study, the resonance trapping process is included in a similar way as that done in other works\(^{52,53}\). The treatment differs depending on whether the pair includes a giant planet or not—the definition of a ‘giant planet’ is given in Supplementary Section 2, whereas an ‘embryo’ refers to a celestial body other than the giant planet.

Case with two embryos. When the orbits of two adjacent embryos (denoted by $i$ and $j$) are converging and get close enough to each other, their separation $(b = |a_i - a_j|)$ impulsively expands at every conjunction. The change in separation via the orbital repulsion is estimated by a linear analysis as

$$\Delta b = \sqrt{\frac{\beta_i}{\beta_j}} \frac{b_i}{\tau_{\nu}},$$

where $\tau_{\nu} = (M_i + M_j)/3 M_{\text{eq}} a_i$ and $a_i = \sqrt{\gamma_p G M_i / \Omega_p}$. The time interval between two conjunctions (that is, the synodic period) is approximately given by

$$\tau_{\text{syn}} \approx \frac{2\pi}{(d\Omega_i/d\Omega_j)} \approx \frac{4 a_i}{3 h \Omega_i^2}.$$  

Therefore, the expansion rate of separation is found to be

$$\frac{db}{dt} \approx \frac{\Delta b}{\tau_{\text{syn}}} \approx \frac{7}{(\tau_{\text{syn}})^{1/4}} a_i b_i.$$  

When the converging speed of the two orbits $\Delta v_{\text{eq}} (\equiv (v_i)_{\text{trap}} - (v_i)_{\text{mig}})$ becomes equal to $db/dt$, the two embryos are assumed to stop approaching, ending up captured in a mean-motion resonance with each other. Therefore, the resultant separation is

$$b_{\text{trap}} = 0.16 \left( \frac{M_i + M_j}{M_{\text{eq}}} \right)^{1/6} \left( \frac{\Delta v_{\text{eq}}}{v_K} \right)^{-1/4} a_i.$$  

Note that we do not specify in which mean-motion resonance these two embryos are really captured, but assume that the mean-motion resonance has a width nearly equal to $b_{\text{trap}}$. In the planetary population synthesis calculations, we calculate $b_{\text{trap}}$ for all the converging pairs. If $b < b_{\text{eq}}$ is satisfied, the pair is treated as captured in the resonance. If $b_{\text{trap}} > 2\sqrt{5} a_i$, the two embryos or planets collide and merge with each other.

After trapped in the resonance, the pair migrates, keeping the ratio of their orbital periods (that is, the ratio of their semi-major axes) unchanged. In that case, their migration rates are different from equation (57). We calculate the loss of angular momenta of the resonantly trapped planets via the interaction between the respective planets and disc and then redistribute the loss to both planets so that the planets migrate with the fixed semi-major axis ratio. The migration rate of the resonantly trapped planet $i$ denoted by $v_{\text{trap,eq}}$ and that of planet $j$ by $v_{\text{trap,trap}}$ (with $a_i/a_j$) such that the angular momentum loss rate, $\dot{L}_{\text{trap}}$, can be expressed as

$$\dot{L}_{\text{trap}} = \frac{d}{dt} \left( M_i \sqrt{G M_i a_i} + M_j \sqrt{G M_j a_j} \right) v_{\text{trap,trap}}.$$  

Here we assume that the eccentricities of both planets are negligible. Thus, $v_{\text{trap,trap}}$ becomes

$$v_{\text{trap,trap}} = \frac{M_i \sqrt{G M_i a_i} + M_j \sqrt{G M_j a_j}}{2a_i}.$$  

The orbital migration terminates at the disc inner edge. The subsequent planets moving inward are trapped in the resonance, and multiple planets line up near the disc inner edge. However, if a heavy enough planet joins such a resonance chain, the heavy planet pushes the planets ahead of itself, particularly the innermost one, into the disc cavity. Following another study\(^{52}\), we adopt the condition for this ‘leakage’ of the planet based on another work\(^{53}\), namely, once the following condition is satisfied, the planet at the disc edge halts the migration of the subsequent planet:

$$\epsilon_i \frac{M_i \mathcal{L}_i}{2 \tau_{\nu} (M_i)} > \sum_{i=1}^{N_{\text{eq}}} \frac{M_i \mathcal{L}_i}{2 \tau_{\nu} (M_i)} > 0,$$  

where $M_i$ is the mass of the $i$th planet ($i = 1$ is the planet at the inner edge), $\mathcal{L}_i \equiv \sqrt{G M_i a_i}$ is the orbital angular momentum, $\tau_{\nu}$ is the migration timescale, and $\epsilon_i$ is the eccentricity damping timescale due to migration, which is given as follows\(^{52}\):

$$\tau_{\nu} = \frac{1}{0.78} \left( \frac{M}{M_*} \right)^{-1} \left( \frac{\Sigma_{\text{d}}^2}{M_*} \right) \left( \frac{b_i}{\Gamma_{\nu}} \right)^{1/3}.$$  

1303
where $e_1$ is the eccentricity of the innermost planet, which is fitted as \( e_1 = 0.02 (\epsilon_{\text{res}}/10)^{1/2} \), (74)

and finally, \( N_{\text{emb}} \) is the number of planets trapped near the disc inner edge. If equation (72) is not satisfied, the innermost planet is pushed into the disc cavity. We assume that the semi-major axis ratio for this pushed-out planet and the next planet is kept constant, until orbital crossing or a giant impact occurs.

**Case including giant planets.** The resonance trapping conditions for a pair of embryo and giant planet and those involving two giant planets are the same as in the two-embryo case. The difference is in the outcome that happens when \( b_{\text{emb}} < 2 \sqrt{3} \eta_{\text{esc}} \). Note that these cases are rare in planetary synthesis models for M dwarfs because giant planets are rarely formed.

For a pair of giant planets, if \( b_{\text{emb}} > 2 \sqrt{3} \eta_{\text{esc}} \), orbital capture occurs. The post-process is calculated following the procedure for dynamical interactions between protoplanets (see step (iii) in the 'Dynamic interaction of multi-body systems and its outcome' section).

For a pair of an embryo and giant planet, the trapping condition is calculated based on another work\(^9\), which gives a criterion for an embryo to enter the feeding zone of a giant planet. This condition is expressed by decreasing the pericenter distance in the separation of two bodies relative to the Hill radius of the giant planet.

The treatment of dynamical interactions differs depending on the number of giant planets in the system: (1) \( N_{\text{giant}} = 0 \), (2) \( N_{\text{giant}} = 1 \), and (3) \( N_{\text{giant}} \geq 2 \). Here we define 'giant planets' as the planets that satisfy both (1) \( M_j > 30 M_E \) and (2) \( e_j > 1 \), where \( e_j = \frac{b_j}{a_j} = \sqrt{2 \eta_{\text{esc}} / \Sigma} \times \frac{\mu_{19, \mu}}{\mu_{30, \mu}} = 1.6 (\frac{M_j}{M_{\odot}})^{1/3} \left( \frac{r}{1 \text{ pc}} \right)^{1/6} \times (\frac{\mu_{19, \mu}}{\mu_{30, \mu}})^{1/2} (78) \)

where \( M_j \) is the jovian mass and \( \mu \) is the mean density of the planet.

(1) If the system has no giant planet or only one giant planet, we first calculate the orbital crossing timescale for every adjacent pair of embryos (\( i \)). The timescale \( t_{\text{cross}} \) follows the fitting formula given as \( \log (t_{\text{cross}} / T_k) = A + B \log (b / 2.3 \eta) \), (79)

where \( T_k \) is the Keplerian period at the semi-major axis of \( a = \sqrt{\pi \Sigma b}, b = (a_j - a_i) \), \( r_{\text{ini}} = (M_j + M_j) / 3 \mu_{\text{ini}}(a_j, a_i) \) and

\[
A = -2.0 + e_i - 0.27 \log \mu + 0.51 \log(\mu) + 0.338 \log(\mu)^2 \tag{80}
\]

\[
B = 18.7 + 1.1 \log \mu - (16.8 + 1.2 \log \mu) e_1 - 0.28 a_i, e_1 = 2 (M_j + M_j) / \mu a_i \tag{81}
\]

Here \( t_i \) (unit: \( \text{yr} \)) is the mean inclination of the two planets. The inclination of a planet in the unit of radians is set to half the eccentricity. The planetary eccentricity before any orbital crossing events is assumed to follow the Rayleigh distribution with a root mean square (r.m.s.) value of \( \sigma = (M_j / 3 \mu_{\text{ini}})^{1/3} \), then after the time interval equal to \( r_{\text{ini}} \), the pair \( (i, j) \) undergoes orbital crossing or sometimes ends up merging. Their resultant masses, semi-major axes and eccentricities are calculated following the procedure presented elsewhere\(^{10}\). They are evaluated so that each of the total mass, orbital energy and Laplace–Runge–Lenz vector is conserved. Finally, if the orbit of the embryo is within 3.5 times the giant planet, the embryo is scattered by the giant planet, and its semi-major axis and eccentricity are modified again\(^{10}\).

(2) If the system has two giant planets (1 and 2) and their separation is \( b = \left| a_i - a_j \right| < 2 \sqrt{3} \eta_{\text{esc}} \), with \( r_{\text{ini}} = (M_j + M_j) / 3 \mu_{\text{ini}}^{1/3} \sqrt{\eta_{\text{esc}}} \), then orbital instability occurs. The orbital elements after instability are calculated following the 'two giants case' discussed in another work\(^{10}\). If orbital instability occurs, all the other embryos are assumed to be ejected from the system. Otherwise, the interactions between embryos are calculated in the same way as that in case (1) above.

(3) If the system has more than two giant planets, the orbital crossing timescale is calculated for every pair of giant planets using equation (79). If any of the derived \( t_{\text{cross}} \) value is larger than the total integration time, the giant planets do not interact with each other, and only the interaction between embryos is calculated as per case (1) above. Otherwise, orbital instability occurs after \( t_{\text{cross}} \), and the resultant masses and orbital elements are derived from the 'three giants case' model discussed elsewhere\(^{11}\). All the other embryos are ejected from the system in this case.

**Initial conditions and parameters.** To start the planetary population synthesis simulations, we perform random samplings of the initial mass \( M_{\text{ini}} \), radius \( r_{\text{ini}} \), metallicity \([\text{Fe/H}]\) and inner edge radius \( r_a \) of the protoplanetary disc gas; the external photoevaporation rate \( \dot{M}_\text{esc} \), and the initial masses and semi-major axes of embryos in the following way. We use the default subroutine \texttt{random_number} in FORTRAN90 to generate random numbers of uniform distribution in \([0, 1]\). To generate random numbers following the normal and Rayleigh distributions, we use the Monte Python method and inverse transform method, respectively.

**Initial conditions for protoplanetary disc.** We determine the initial properties of the protoplanetary disc by scaling recent observation results for stars of \( \sim 1 \, \text{M}_\odot \). We adopt the fitting formula for the disc mass for \( 1 \, \text{M}_\odot \) stars, the log-normal distribution with the mean of \( \log(M/1 \, \text{M}_\odot) = 1.49 \) and the standard deviation \( \sigma = 0.35 \), which were derived in another work\(^{12}\) derived from the observational results discussed elsewhere\(^{13}\). The minimum and maximum disc masses are set to \( 4 \times 10^{-3} \, \text{M}_\odot \) and \( 0.16 \, \text{M}_\odot \), which roughly correspond to the lightest and heaviest samples in ref. \(^{13}\), respectively. We also assume that the mean, minimum and maximum disc masses are proportional to the stellar masses\(^{13}\). The disc gas radius is calculated as

\[
disc = 10 \left( \frac{M_{\text{ini}}}{2 \times 10^{-3} \, \text{M}_\odot} \right)^{0.625} \text{AU}, \tag{83}
\]

which is taken from the observational trend derived in another work\(^{14}\).

The disc metallicity \([\text{Fe/H}]\) follows the normal distribution with \( \mu = -0.02 \) and \( \sigma = 0.22 \) (ref. \(^{12}\)). The range of values is limited to \(-0.6 < \text{[Fe/H]} < 0.5\). We use the same distribution regardless of the stellar mass.

We assume that the disc inner edge is located at the co-rotation radius where the Kepler rotation period is equal to the rotation period of the central star. Here the stellar rotation period is assumed to follow the log-normal distribution with \( \log(p/\text{days}) = 0.676 \) and \( \sigma = 0.306 \) based on the observational results of young stellar objects\(^{15}\). The minimum of \( r_a \) is set to be the initial stellar radius \( R \). We also use the same distribution for all stellar types.

The external photoevaporation rate \( \dot{M}_{\text{esc}} \) generally depends on the population of nearby massive stars. Following another work\(^{16}\), we set the distribution of \( M_{\text{esc}} \) so that the mean value of the resultant disc lifetime is located at \( \sim 3 \, \text{Myr} \) (refs. \(^{12,17}\)) and that it has a deviation of about half an order, regardless of the stellar mass. Since the disc gas radius \( r_a \) is determined only by the disc mass \( M_{\text{esc}} \), the disc lifetime depends on \( M_{\text{esc}} \) and \( M_{\text{esc}} \) for a given \( r_a \). Then, we find that for a star of \( 0.3 \, \text{M}_\odot \), and \( e_1 = 2 \times 10^{-3} \), the log-normal distribution with \( \log(\mu/\text{yr}) = -6.0 \) and \( \sigma = 0.5 \) accounts for the above distribution. Also, the stellar mass dependence of \( M_{\text{esc}} \propto M_{\odot}^{1.4} \) is found to be suitable for the star in the range of \( 0.1 \, \text{M}_\odot \leq M \leq 0.5 \, \text{M}_\odot \).
Table 1 | Parameters used in calculations

| Symbol | Meaning | Value | Primarily used equation |
|--------|---------|-------|-------------------------|
| $\alpha_{\text{esc}}$ | Parameter for effective turbulent viscosity | $2.0 \times 10^{-3}$ | (3) |
| $\alpha_{\text{vis}}$ | Parameter for turbulent viscosity | $2.0 \times 10^{-4}$ | (44) |
| $N_{\text{grid,esc}}$ | Number of grids for gas disc | 500 | |
| $N_{\text{grid,solid}}$ | Number of grids for solid disc | 1,000 | |
| $r_{\text{esc}}$ | Outer boundary radius for gas disc | 1,000 AU | (5) |
| $r_{\text{solid}}$ | Solid disc radius | 0.5 AU | (17) |
| $\Phi$ | Ionizing XUV photon luminosity | $1.0 \times 10^{40}$ $\text{g}^{-1}$ | (7) |
| $m_{\text{p}}$ | Planetesimal mass | $1.0 \times 10^{15}$ g | (27) |
| $\rho_{\text{p}}$ | Planetesimal material density | $3.0 \times 10^{-3}$ g cm$^{-3}$ | (27) |
| $C_{\text{v,sh}}$ | Specific heat of solid core for constant volume | $1.2 \times 10^{9}$ erg g$^{-1}$ K$^{-1}$ | (35) |
| $K_i$ | Typical value of $K$ for which the co-rotation torque becomes ineffective | 20.0 | (64) |

Initial conditions for planetary embryos. Initially, 50 planetary embryos with a mass of 0.01$M_\oplus$ are placed log-uniformly from $r_{\text{esc}}$ to $r_{\text{solid}}$. Here the initial separations of all the adjacent embryos are larger than the feeding zone width (10$r_{\text{esc}}$) for the local isolation mass $M_{\text{esc}}$ given by

$$M_{\text{esc}} = 0.16 \left( \frac{\Sigma_{\text{gas}}}{10 \text{ g cm}^{-2}} \right)^{1/3} \left( \frac{a}{1\text{AU}} \right)^{1/4} \left( \frac{M_\oplus}{M_{\text{esc}}} \right)^{1/2} \left( \frac{L_{\text{esc}}}{L_\oplus} \right)^{1/2}. \quad (84)$$

Therefore, in fairly massive solid discs, the initial number of planetary embryos can be smaller than 50.

Input parameters. The parameters and their fiducial values are summarized in Table 1. These values are used in our calculations, unless mentioned otherwise.

Probability density. To show the distribution of water mass fraction in the synthesized planets (Figs. 2 and 3), we use the probability densities calculated by

$$P(D) = \frac{N_i}{N_{\text{tot}} \Delta \log(M_{\text{esc}}/M_{\text{esc}})} \quad (85)$$

where $N_i$ is the number of planets in the $i$th bin, $N_{\text{tot}} = \sum N_i$ is the total number of planets and $\Delta \log(M_{\text{esc}}/M_{\text{esc}})$ is the bin width.

Data availability. All data from the simulation are available via GitHub at https://github.com/TadahiroKimura/Kimura-Ikoma2022. Source data are provided with this paper.

Code availability. The numerical code used in the current study is available from the corresponding authors on reasonable request.
98. Tanaka, H. & Ward, W. R. Three-dimensional interaction between a planet and an isothermal gaseous disk. II. Eccentricity waves and bending waves. *Astrophys. J.* **602**, 388 (2004).

99. Shiraishi, M. & Ida, S. Infall of planetesimals onto growing giant planets: onset of runaway gas accretion and metallicity of their gas envelopes. *Astrophys. J.* **616** (2004).

100. Artymowicz, P. Disk-satellite interaction via density waves and the eccentricity evolution of bodies embedded in disks. *Astrophys. J.* **419**, 166 (1993).

101. Iwasaki, K., Emori, H., Nakazawa, K. & Tanaka, H. Orbital stability of a protoplanet system under a drag force proportional to the random velocity. *Publ. Astron. Soc. Jpn.* **54**, 471–479 (2002).

102. Zhou, J.-L., Lin, D. N. C. & Sun, Y.-S. Post-oligarchic evolution of protoplanetary embryos and the stability of planetary systems. *Astrophys. J.* **666**, 423 (2007).

103. Emsenhuber, A. The New Generation Planetary Population Synthesis (NGPPS). II. Planetary population of solar-like stars and overview of statistical results. *Astron. Astrophys.* **656**, A70 (2021).

104. Tychoniec, Ł. The VLA Nascent Disk and Multiplicity Survey of Perseus Protostars (VANDAM). IV. Free–free emission from protostars: links to infrared properties, outflow tracers, and protostellar disk masses. *Astrophys. J.* **238**, 19 (2018).

105. Andrews, S. M., Rosenfeld, K. A., Kraus, A. L. & Wilner, D. J. The mass dependence between protoplanetary disks and their stellar hosts. *Astrophys. J.* **771**, 129 (2013).

106. Santos, N. C. Spectroscopic metallicities for planet-host stars: extending the samples. *Astron. Astrophys.* **437**, 1127–1133 (2005).

107. Venuti, L. CSI 2264: investigating rotation and its connection with disk accretion in the young open cluster NGC 2264. *Astron. Astrophys.* **599**, A23 (2017).

108. Burn, R. The New Generation Planetary Population Synthesis (NGPPS). IV. Planetary systems around low-mass stars. *Astron. Astrophys.* **656**, A72 (2021).

109. Mamajek, E. E. Initial conditions of planet formation: lifetimes of primordial disks. *AIP Conf. Proc.* **1158**, 3 (2009).

**Acknowledgements**

This work is supported by JSPS KAKENHI (nos. JP18H05439, JP21H01141 and JP22J11725). T.K. is a JSPS Research Fellow, and also supported by the International Graduate Program for Excellence in Earth-Space Science (IGPEES).

**Author contributions**

Both authors contributed equally to this work. M.I. conceived the original idea and supervised this project. T.K. developed the entire model of planetary population synthesis partly using a few modules that M.I. had developed. T.K. carried out the numerical simulations and analysed the simulation results. Both authors discussed the results and implications and wrote the paper.

**Competing interests**

The authors declare no competing interests.

**Additional information**

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41550-022-01781-1.

Correspondence and requests for materials should be addressed to Tadahiro Kimura or Masahiro Ikoma.

Peer review information *Nature Astronomy* thanks Nicolas Cowan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2022.