

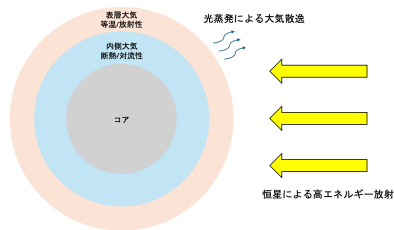
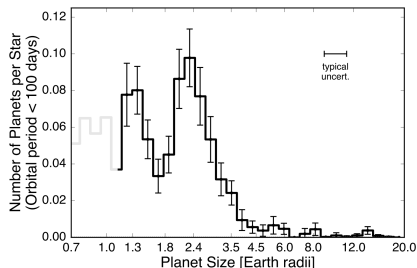
1. 系外惑星の半径分布に現れる「蒸発の谷」の解析的導出
2. 物理モデルを必要としない逆問題の解法：ニューラルネットワークと定性的正則化を用いた方法

三田村彰大

# Owen & Wu 2017; The Evaporation Valley in the Kepler Planets

1/8

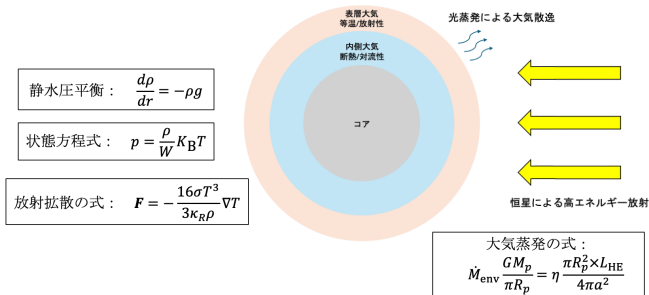
- ・系外惑星の半径分布には特徴的な谷がある
- ・恒星放射による大気蒸発モデルによって**数值的**に予言されていた。
- ・しかし、**解析的**な導出はこの論文以前には存在していない



第一原理から惑星大気の時間発展を記述する量を導出し、谷の存在を解析的に導く。

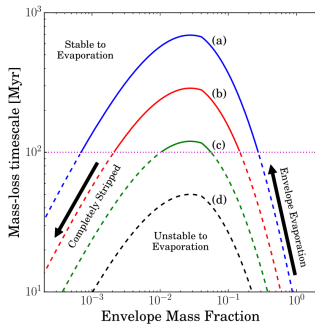
$$\text{大気の質量減少のタイムスケール } \tau = -\frac{X}{\dot{X}} \quad \left( X: \text{惑星大気のコア質量比} \right)$$

→  $\tau$  を第一原理から導出する



大気蒸発に伴う大気の質量減少のタイムスケール  $\tau$  は次のように計算される。

$$\tau = \left\{ \begin{array}{l} 210\text{Myr} \times \left(\frac{\eta}{0.1}\right)^{-1} \left(\frac{L_{\text{HE}}}{10^{-3.5}L_{\odot}}\right)^{-1} \\ \times \left(\frac{P}{10\text{days}}\right)^{1.41} \left(\frac{M_*}{M_{\oplus}}\right)^{0.52} \left(\frac{f}{1.2}\right)^{-3} \\ \times \left(\frac{\tau_{\text{KH}}}{100\text{Myr}}\right)^{0.37} \left(\frac{\rho M_{\oplus}}{5.5\text{g cm}^{-3}}\right)^{0.18} \left(\frac{M_c}{5M_{\oplus}}\right)^{1.42} \end{array} \right\} \times \left\{ \begin{array}{l} \left(\frac{R_{\text{atm}}}{R_{\text{core}}}\right)^{1.57} \quad \left(R_{\text{atm}}/R_{\text{core}} \ll 1\right) \\ \text{薄い大気} \\ \left(\frac{R_{\text{atm}}}{R_{\text{core}}}\right)^{-1.69} \quad \left(R_{\text{atm}}/R_{\text{core}} \gg 1\right) \\ \text{厚い大気} \end{array} \right.$$



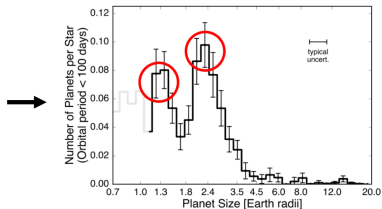
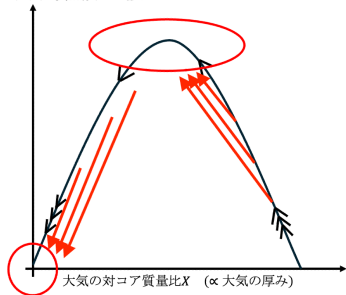
**Figure:** 横軸は  $X$  ( $\propto$  惑星大気の厚み)、縦軸は  $\tau$  (= 大気の質量減少のタイムスケール)。(a) $\rightarrow$ (d) は恒星に近くなる向きを表す。

大気が分厚い時と薄い時で大気は激しく萎むが、中程度の厚みを持つ大気はあまり萎まない。

- 厚い大気を持つ惑星たちは蒸発により  $R_{\text{大気}} \approx R_{\text{コア}}$  あたりに溜まっていく
- 薄い大気を持つ惑星たちは蒸発により大気を完全に剥ぎ取られ  $R_{\text{大気}} \approx 0$  に溜まっていく

→ 二集団性（二峰性）を解析的に導けた！

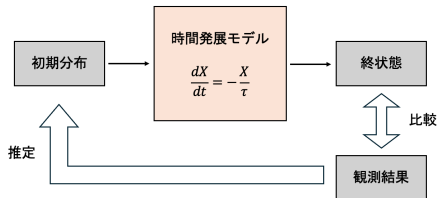
$$\tau = -\frac{X}{\dot{X}} \left( \propto \frac{1}{\text{大気}の\text{質量}減少\text{の}速度} \right)$$



分かったこと：

系外惑星の半径分布に現れる「蒸発の谷」は、蒸発による惑星大気質量減少のタイムスケールが  $R_{\text{atm}} = R_{\text{core}}$  となる位置で最大となり、その両側で減衰することに起因する。

また、時間発展モデルを使うと逆問題的に初期条件を推定できる。

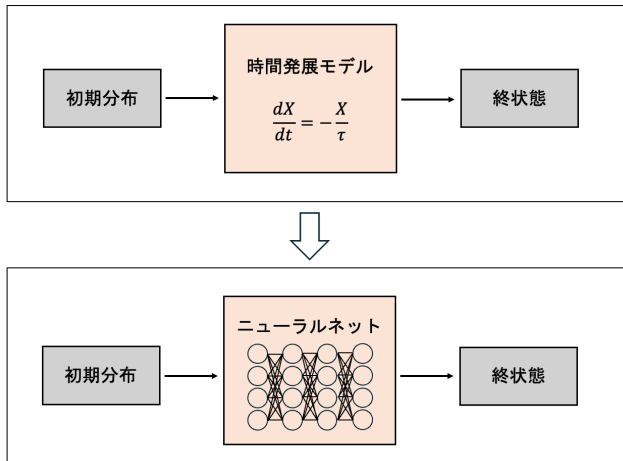


- 初期状態で  $X \sim \mathcal{O}(1)\%$  程度の大气を持っていないといけない。
  - → 惑星の大气獲得は原始惑星系円盤から行われなければならない。
- 観測を再現するためには惑星のコア組成は iron-silicate でないといけない。
  - → アイスラインの外からの immigration モデルは否定される。

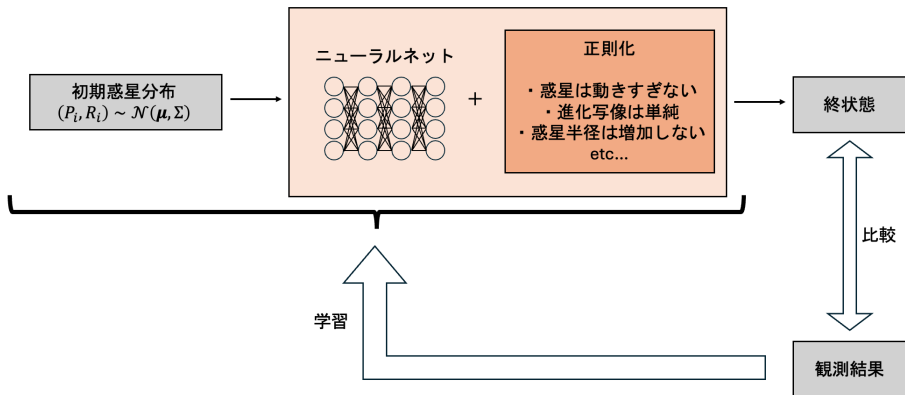
# 物理モデルを必要としない逆問題の解法： NN と定性的正則化を用いた方法

6/8

解析モデルを構築できていない状態で逆問題を解くことはできないか？



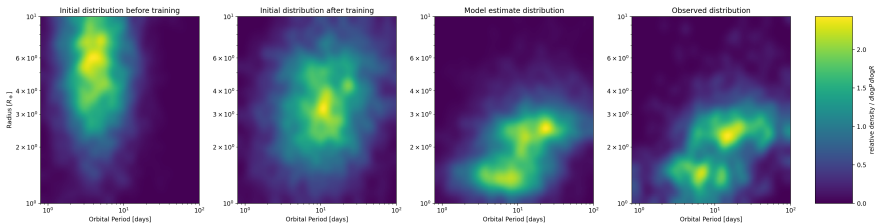
- ・初期惑星分布とニューラルネットをまとめて学習する。
- ・ニューラルネットは柔軟すぎて初期惑星分布の動きを吸収してしまう。
  - → 物理モデルに関する直感的な情報を正則化として入れる。



分かったこと：

- どんなに正則化を強くしても逆問題の解が学習初期条件の影響を受けてバラついてしまう問題を完全に無くすることができず、手法としての有効性は低いと思われる。
- ただし、正則化の強さ、学習初期条件のばらつきによらず惑星初期分布の周期方向の中心が 10days に強く収束したのは特筆すべき。

※ 情報量基準の定め方の恣意性など、手法としての問題性は様々あると思われる。



**Figure:** 左から：学習前の初期惑星分布、学習後の初期惑星分布、学習後の初期惑星分布を学習後の NN で進化させて得られる惑星分布、観測で得られている初期惑星分布。

※ 以下は詳細解説スライド

# 序論：Kepler 宇宙望遠鏡による太陽系外惑星観測

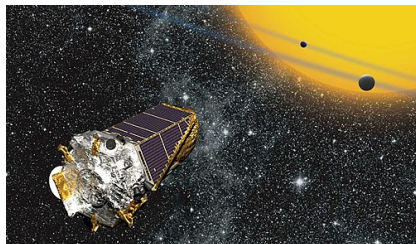
ケプラー宇宙望遠鏡（2009）による太陽系外惑星観測：約 2800 個の太陽系外惑星を発見。（kepler planets/kepler 惑星）

→トランジット法による観測：

- 惑星が恒星の前を通過する際の減光度合い → 惑星の半径  $R$

$$\text{減光率 } \Delta F = (R_p/R_*)^2$$

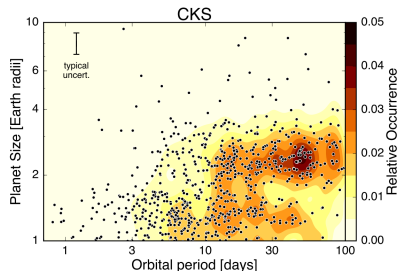
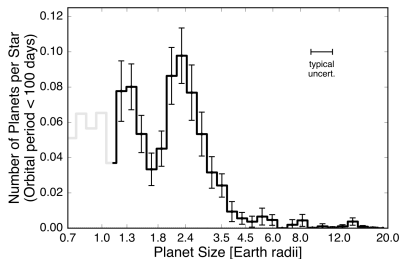
- 惑星が恒星の前を通過する周期 → 惑星の周期  $P$



# 序論：Kepler 惑星の半径分布に現れる「蒸発の谷」

California Kepler survey (2017) による Kepler 惑星の解析により、次のことがわかった。

- $< 4R_{\oplus}$  程度の近接・小型惑星が多く見つかった。
- 近接・小型惑星の半径分布には特徴的な「谷」がある。
  - $\sim 1.8R_{\oplus}$  ( $R_{\oplus}$  は地球半径) の惑星の欠乏
  - $< 1.3R_{\oplus}$  と  $2.6R_{\oplus}$  の大きさを持つ惑星が多数存在。

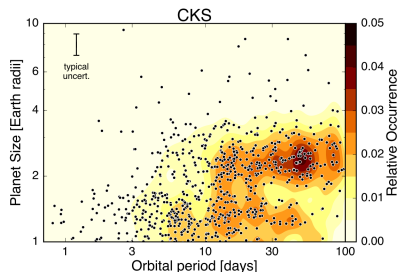
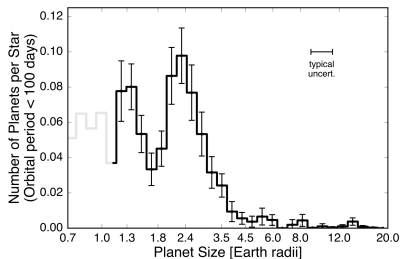


# 序論：Owen & Wu 2017 以前の研究

恒星由来の極端紫外線によって惑星の大気が剥ぎ取られる（蒸発する）ことにより惑星半径が変化し、特徴的な半径の谷が現れるのではないか？

- 蒸発による「谷」の再現は数値モデルにおいて成功していた。
- → 解析モデル（理論モデル）は存在していない。

「蒸発の谷」を再現するような解析モデル（第一原理モデル）を構築しよう！



# 何を計算したいのか

最終的には、第一原理から惑星の大気進化モデルを構築し、惑星半径に「谷」が現れる様子を再現したい。

→ 惑星半径（より厳密には惑星大気 Envelope）の時間発展の式を導出して、初期条件を定めて時間発展させた時に「谷」が現れるかを見れば良い！

惑星大気 Envelope のコアに対する質量比  $X = \frac{M_{\text{env}}}{M_{\text{core}}}$  と、惑星大気 Envelope 質量の減少のタイムスケール  $\tau = \frac{X}{\dot{X}}$  を用いて時間発展は次のような微分方程式で記述できる。

$$\frac{dX}{dt} = -\frac{X}{\tau}$$

→ よって、 $\tau = \frac{X}{\dot{X}}$  と  $X = \frac{M_{\text{env}}}{M_c}$  を第一原理から導出できればいい！

Owen&Wu の論文がやっていること：

- 適切な解析モデルを設定し、第一原理から惑星大気 Envelope の時間発展の式を導出する。
- 適切に初期条件を定め、惑星集団を時間発展させた時に惑星の半径分布に「谷」が現れるかどうかを確認する。

# 解析の準備 1 : 仮定するモデル

固体コアと理想気体状態方程式に従うガス envelope からなる惑星を考える。

- Envelope の内側は断熱的な温度プロファイル+対流による熱輸送
- Envelope の外側は等温+放射による熱輸送

惑星は恒星由来の高エネルギー放射による光蒸発によって大気を剥ぎ取られ、大きさが変化する。

- 蒸発効率が恒星からの放射の強さで決まるような蒸発モデル (Energy-Limited モデル) を考える。

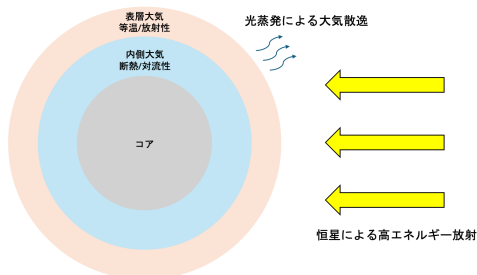


Figure: 解析モデル。論文を元に作成。

# 解析の準備 2 : 計算で用いる第一原理

惑星および惑星大気が従う次の物理法則群が解析モデルの出発点となる。

静水圧平衡 :

$$\frac{d\rho}{dr} = -\rho g$$

状態方程式 :

$$p = \frac{\rho}{W} K_B T$$

断熱気体に関する式 :

$$p = K\rho^\gamma$$

放射拡散の式 :

$$\mathbf{F} = -\frac{16\sigma T^3}{3\kappa_R \rho} \nabla T$$

大気蒸発を記述する式 :

$$\dot{M}_{\text{env}} \frac{GM_p}{\pi R_p} = \eta \frac{\pi R_p^2 \times L_{\text{HE}}}{4\pi a^2}$$

ケプラーの第三法則 (円軌道) :

$$\frac{p^2}{a^3} = \frac{4\pi}{GM_*}$$

# 解析の準備 3 : 計算途中で用いる近似

今回のモデルはあくまで簡易的なモデルであり、計算の途中で多数の近似を用いている。

近似の例：

- 惑星大気は理想気体方程式に従う。
- 惑星大気の組成は原始太陽組成に近いと考える。
- 惑星大気 envelope はコアに比べて十分軽く、その重力への寄与は無視できる。
- 惑星大気は重力に強く捕捉されており、脱出速度は音速に比べ十分小さい。
- 大気散逸は恒星放射によるもののみとし、恒星の重力による散逸などは考えない。
- **恒星の高エネルギー放射による蒸発の効率是一定とする。**
- 惑星の光度においてコア光度の寄与は考えない。
- 知られている近似式や経験式を用いる。

等温層の厚み：  $R_p - R_{\text{rcb}} \approx 6H$

不透明度：  $\kappa_R \approx \kappa_0 P^\alpha T^\beta$  など

恒星の放射輝度：  $L_*/L_\oplus = (M_*/M_\odot)^{3.2}$

# 解析結果 1 : 時間発展モデル

我々が今やろうとしていることは「蒸発の谷」を再現するような物理モデルの構築であり、それは惑星大気の時間発展を記述する微分方程式として表されるのであった。

$$\frac{dX}{dt} = -\frac{X}{\tau} \quad \left( \begin{array}{l} X: \text{大気対コア質量比} \\ \tau: \text{大気の質量減少のタイムスケール} \end{array} \right)$$

よって  $\tau$  と  $X$  を第一原理から計算すればよく、その結果は次のようになる。

$$\tau \propto \begin{cases} \Delta R^{1.57} & (\Delta R/R_c \ll 1, \text{薄い大気}) \\ \Delta R^{-1.69} & (\Delta R/R_c \gg 1, \text{厚い大気}) \end{cases}$$

$$X \propto \begin{cases} \Delta R^{1.57} & (\Delta R/R_c \ll 1, \text{薄い大気}) \\ \Delta R^{1.31} & (\Delta R/R_c \gg 1, \text{厚い大気}) \end{cases}$$

$\Delta R$  : 大気の厚み,  $R_c$  : コア半径

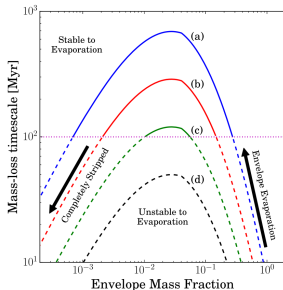
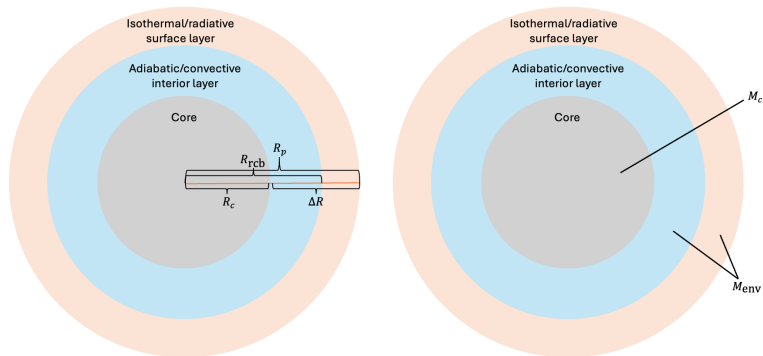


Figure:  $\tau$ ,  $X$  プロット。横軸は対コア質量比  $X$ 、縦軸は大気の質量減少のタイムスケール  $\tau$ 。Fulton et al. より抜粋。

# 計算の概略 1 : ノーテーション



$R_{\oplus}, R_{\odot}$  : 地球半径, 太陽半径、 $M_{\oplus}, M_{\odot}$  : 地球質量, 太陽質量、 $T_{eq}$  : 等温層の温度、  
 $\rho_{rcb}$  : 境界層での密度、 $\mu$  : 平均分子量

## 計算の概略2 : X の導出

$X = M_{\text{env}}/M_c$  なので、 $M_{\text{env}}$  を計算すれば  $X$  は求まる。今、静水圧平衡  $dp/dr = -\rho g$  及び断熱気体の式  $p = K\rho^\gamma$  を用いると

$$\rho(r) = \rho_{\text{rcb}} \left[ 1 + \frac{\gamma - 1}{\gamma} \left( \frac{GM_c}{\rho_{\text{rcb}}^{\gamma-1} K R_{\text{rcb}}} \right) \left( \frac{R_{\text{rcb}}}{r} - 1 \right) \right]^{\frac{1}{\gamma-1}}$$

のようにして大気の密度が求まるので、大気質量は  $M_{\text{env}} = \int_{R_c}^{R_p} 4\pi r^2 \rho dr$  より

$$M_{\text{env}} = 4\pi R_p^3 \rho_{\text{rcb}} \left( \nabla_{\text{ad}} \frac{GM_c}{c_s^2 R_p} \right)^{\frac{1}{\gamma-1}} I_2$$

となる。ただし

- $c_s$  は境界層での音速  $c_s^2 = p/\rho$
- $\nabla_{\text{ad}} = (\gamma - 1)/\gamma$
- $I_n = \int_{R_c/R_p}^1 x^n (x^{-1} - 1)^{1/(\gamma-1)} dx$

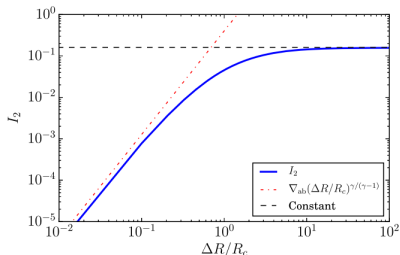
であり、また途中で

- 重力による束縛が十分強い ( $v_{\text{esc}}^2 \sim \frac{GM_c}{R_p} \gg c_s^2$ )

という近似を用いている。

## 計算の概略2 : X の導出 (続き) : 解の分裂

X の式中に現れる積分は  $I_2 \sim \nabla_{\text{ad}} \left( \frac{\Delta R}{R_c} \right)^{\frac{\gamma}{\gamma-1}}$  for  $\Delta R/R_c \ll 1$  and  $I_2 \sim \text{const.}$  for  $\Delta R/R_c \gg 1$  と近似できる。(下図参照。)



これにより大気が薄い場合 ( $\Delta R/R_c \ll 1$ ) と大気が厚い場合 ( $\Delta R/R_c \gg 1$ ) で解が分裂する。

$$X \approx \left[ 3 (\nabla_{\text{ad}})^{\frac{\gamma}{\gamma-1}} \right] \frac{\rho_{\text{rcb}}}{\rho_{\text{core}}} \left( \frac{GM_c}{c_s^2 R_c} \right)^{\frac{1}{\gamma-1}} \left( \frac{\Delta R}{R_c} \right)^{\frac{\gamma}{\gamma-1}} \quad (\Delta R/R_c \ll 1)$$

$$X \approx \left[ 3 C_{I_2^\infty} (\nabla_{\text{ad}})^{\frac{1}{\gamma-1}} \right] \frac{\rho_{\text{rcb}}}{\rho_{\text{core}}} \left( \frac{GM_c}{c_s^2 R_c} \right)^{\frac{1}{\gamma-1}} \left( \frac{\Delta R}{R_c} \right)^{\frac{3\gamma-4}{\gamma-1}} \quad (\Delta R/R_c \gg 1)$$

## 計算の概略2 : X の計算 (続き)

X の式中に  $\rho_{\text{rcb}}$  は放射拡散の式  $F = -\frac{16\sigma T^3}{3\kappa_R\rho}\nabla T$  及びロスランド不透明度の近似式  $\kappa_R \approx \kappa_0 P^\alpha T^\beta$  と、ケルビン・ヘルムホルツタイムスケール  $\tau_{\text{KH}} = U_{\text{potential}}/(\text{Luminosity of planet})$  を用いることで次のように評価できる。

$$\rho_{\text{rcb}} = \left(m_H \nabla_{ab}^{\frac{1}{1+\alpha}}\right) \left(\frac{\mu}{k_B}\right) \left[\left(\frac{l_2}{l_1}\right) \frac{64\pi\sigma T_{\text{eq}}^{3-\alpha-\beta} R_p \tau_{\text{KH}}}{3\kappa_0 M_c X}\right]^{1/(1+\alpha)}$$

よってこれを代入すれば最終的な X の式を得ることができる。

$$X \approx C_{\text{small}} \times \left(\frac{l_1}{l_2}\right)^{n_l} \mu^{n_\mu} \kappa_0^{n_\kappa} T_{\text{eq}}^{n_T} \tau_{\text{KH}}^{n_\tau} \rho_\oplus^{n_\rho} M_c^{n_M} \left(\frac{\Delta R}{R_c}\right)^{n_a} \quad (\Delta R/R_c \ll 1) \quad (1)$$

$$X \approx C_{\text{large}} \times \left(\frac{l_1}{l_2}\right)^{n_l} \mu^{n_\mu} \kappa_0^{n_\kappa} T_{\text{eq}}^{n_T} \tau_{\text{KH}}^{n_\tau} \rho_\oplus^{n_\rho} M_c^{n_M} \left(\frac{\Delta R}{R_c}\right)^{n_b} \quad (\Delta R/R_c \gg 1) \quad (2)$$

ただし  $C_{\text{small}}, C_{\text{large}}$  は定数であり、指数  $n_i$  は次ページに示す通り。これを用いると

$$X \propto \begin{cases} \Delta R^{1.57} & (\Delta R/R_c \ll 1, \text{薄い大気}) \\ \Delta R^{1.31} & (\Delta R/R_c \gg 1, \text{厚い大気}) \end{cases}$$

として先に示した式を得る。

## 計算の概略2：Xの計算（続き）

Xの表式に現れる各係数の指数について：

- ロスランド不透明度の近似  $\kappa_R \approx \kappa_0 P^\alpha T^\beta$  について、定質量・恒星放射下にある H/He envelope に関する結果を使う。 $(\alpha = 0.68, \beta = 0.45)$
- 断熱気体に関する式  $p = K\rho^\gamma$  に現れる  $\gamma$  を  $\gamma = 5/3$  と考える。

この時、

$$n_a = \frac{\gamma}{\gamma-1} \frac{\alpha+1}{\alpha+2} \approx 1.57 \quad n_b = \left( \frac{3\gamma-4}{\gamma-1} + \frac{1}{\alpha+1} \right) \frac{\alpha+1}{\alpha+2} \approx 1.31$$

$$n_l = n_\tau = \frac{1}{\alpha+2} \approx 0.37 \quad n_\mu = \left( 1 + \frac{1}{\gamma-1} \right) \frac{\alpha+1}{\alpha+2} \approx 1.57 \quad n_\kappa = -\frac{1}{\alpha+2} \approx -0.37$$

$$n_T = \left( \frac{3-\alpha-\beta}{\alpha+1} - \frac{1}{\gamma-1} \right) \frac{\alpha+1}{\alpha+2} \approx -0.24$$

$$n_\rho = \left( \frac{1}{3} \left( \frac{1}{\gamma-1} - \frac{1}{\alpha+1} \right) - 1 \right) \frac{\alpha+1}{\alpha+2} \approx -0.44 : \text{論文中ではこの値に誤植があると思われる}$$

$$n_M = \left( \frac{2}{3} \left( \frac{1}{\gamma-1} - \frac{1}{\alpha+1} \right) \right) \frac{\alpha+1}{\alpha+2} + \frac{n_\rho}{4} \approx 0.17$$

# 計算の概略3： $\tau$ の計算

$\tau = -X/\dot{X}$ なので、 $\dot{X}$ が求まれば先の結果と合わせて $\tau$ が求まるが、 $\dot{X} = \dot{M}_{\text{env}}/M_c$ であり、大気蒸発を記述する式より  $\dot{M}_{\text{env}} \frac{GM_p}{\pi R_p} = \eta \frac{\pi R_p^2 \times L_{\text{HE}}}{4\pi a^2}$  とわかっているのでこれは求めることができ次のようになる。

$$\tau = \left\{ \begin{array}{l} 210\text{Myr} \times \left(\frac{\eta}{0.1}\right)^{-1} \left(\frac{L_{\text{HE}}}{10^{-3.5}L_{\odot}}\right)^{-1} \\ \times \left(\frac{P}{10\text{days}}\right)^{1.41} \left(\frac{M_*}{M_{\oplus}}\right)^{0.52} \left(\frac{f}{1.2}\right)^{-3} \\ \times \left(\frac{\tau_{\text{KH}}}{100\text{Myr}}\right)^{0.37} \left(\frac{\rho_{M_{\oplus}}}{5.5\text{g cm}^{-3}}\right)^{0.18} \left(\frac{M_c}{5M_{\oplus}}\right)^{1.42} \end{array} \right\} \times \left\{ \begin{array}{l} \left(\frac{\Delta R}{R_c}\right)^{1.57} \quad (\Delta R/R_c \ll 1) \\ \left(\frac{\Delta R}{R_c}\right)^{-1.69} \quad (\Delta R/R_c \gg 1) \end{array} \right.$$

ただし途中で

- stellar tides は無視 (恒星の重力による大気の剥ぎ取りはない。)
- 恒星の輝度に関する経験則  $L_*/L_{\odot} = (M_*/M_{\odot})^{3.2}$

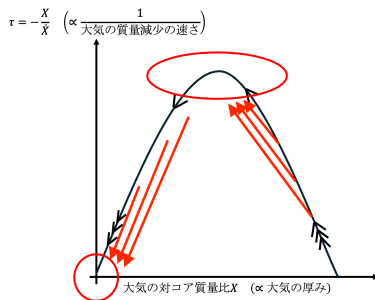
という近似を使っている。

# 解析結果 1 : 時間発展モデル

$\tau$  を  $X$  に対してプロットすると、大気の質量減少のタイムスケール  $\tau$  は惑星半径がコア半径の 2 倍となる位置で最大となりその両側で減衰するということがわかる。

→ タイムスケールは大気が萎むスピードに反比例しているため、大気が分厚い時と薄い時で大気は激しく萎むが、中程度の厚みを持つ大気はあまり萎まない。

- 厚い大気を持つ惑星たちは蒸発によりグラフのピークあたりに溜まっていく。
- 薄い大気を持つ惑星たちは蒸発により大気を完全に剥ぎ取られグラフの左下に溜まっていく。



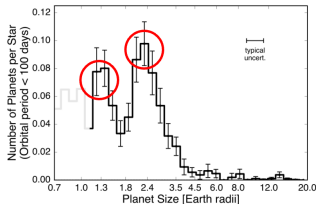
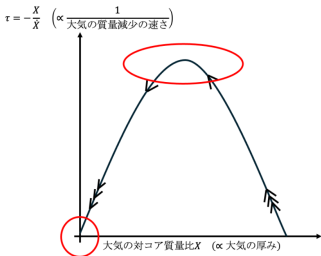
# 解析結果 1 : 時間発展モデル

よって質量減少のタイムスケール  $\tau$  の  $X$  依存性を表すグラフの形を観察することで、恒星放射による大気蒸発によって惑星集団は

- 蒸発により完全に大気を剥ぎ取られてしまうような、比較的薄い大気を持つ集団
- コア質量比  $X$  が同じ値に収束するような、比較的厚い大気を持つ集団

に分けられることがわかる。この二集団性が「蒸発の谷」において二峰性が現れることと対応している。

→ 「蒸発の谷」を生む解析モデルを第一原理から導けた！



# 解析結果 1 : 時間発展モデル

$\tau$  のグラフが上記のような形になる物理的な背景について述べておく。

$$\tau = -\frac{X}{\dot{X}} \propto \frac{(\text{大気の厚み/コア半径})^{1.3 \sim 1.5}}{\text{大気蒸発のスピード}}$$

であり、 $\dot{M}_{\text{env}} \frac{GM_p}{\pi R_p} = \eta \frac{\pi R_p^2 \times L_{\text{HE}}}{4\pi a^2}$  なので (大気蒸発のスピード)  $\propto R_p^3$  である。よって、

- 大気が分厚くなるほど大気蒸発のスピードが早くなる。(直感的には恒星からの高エネルギー放射を受ける面積が増える。)
- → そのスピードは大気の質量の増加スピードより指数的に速いので、タイムスケールは大気が厚いほど小さくなっていく。

また逆に、大気が薄くなると  $\Delta R \ll R_c$  より  $R_p = R_c + \Delta R \sim R_c$  であるから、

- 大気が薄くなると大気蒸発のスピードはほぼ一定となる。(  $R_p \sim R_c \sim \text{const.}$  )
- しかし大気が薄くなるほど大気の質量は減少し続ける。(  $\Delta R/R_c \neq \text{const.}$  )
- → 大気が薄いほどタイムスケールは小さくなっていく。

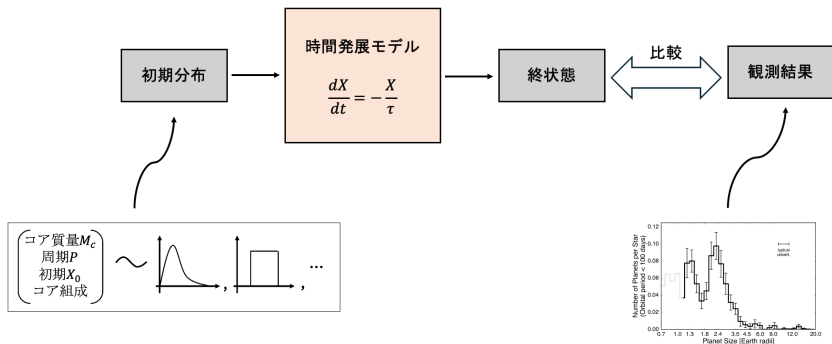
よって  $\tau$  の山なりの構造が出てくる。

# 解析結果 2 : 初期惑星集団を仮定した時間積分

先の考察はあくまで定性的に「蒸発の谷」を導けたに過ぎない。

→ 初期惑星集団を用意して時間積分を行い、谷が再現できるか確認する。

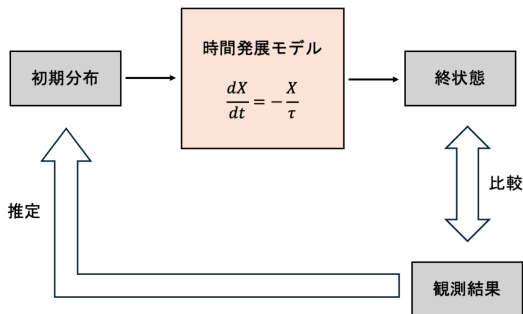
ただし、時間発展させる前に初期惑星集団の具体的な形状（コア質量の分布、恒星からの距離の分布、etc.）を決定しなければならない。



## 解析結果 2 : 初期惑星集団を仮定した時間積分

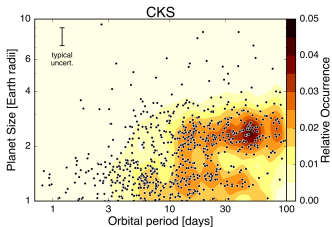
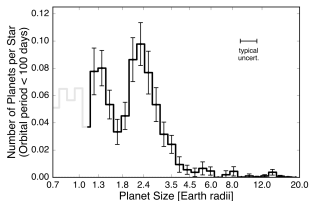
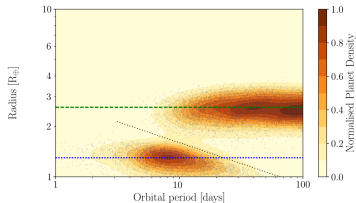
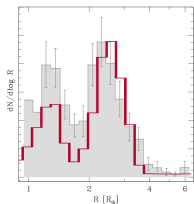
論文においては、「進化させた時に谷を再現するような初期惑星分布」を採用している。すなわち、構築した解析モデルを用いて逆問題的に初期分布を推定している。

→ この点は循環論法になってしまっている可能性がある。



# 解析結果 2 : 初期惑星集団を仮定した時間積分

時間積分の結果：(上：解析モデル、下：観測)



→ 「蒸発の谷」をピーク位置・谷の位置・ $R_p$  大の欠損まで含めて非常に良く再現できている！また周期依存性のおおよその傾向もつかめている！

# 結果の考察

今回は、惑星の大気進化の簡易的な解析モデルを構築した。その結果として、

- 「蒸発の谷」を再現できた。
- 「蒸発の谷」の成因は、惑星半径がコア半径の2倍となる位置でタイムスケールが最大となり、その両側でタイムスケールが減衰することにある。
- 「蒸発効率は一定」という近似が一番影響が大きく、この近似を外すとさらによく観測を再現できる。

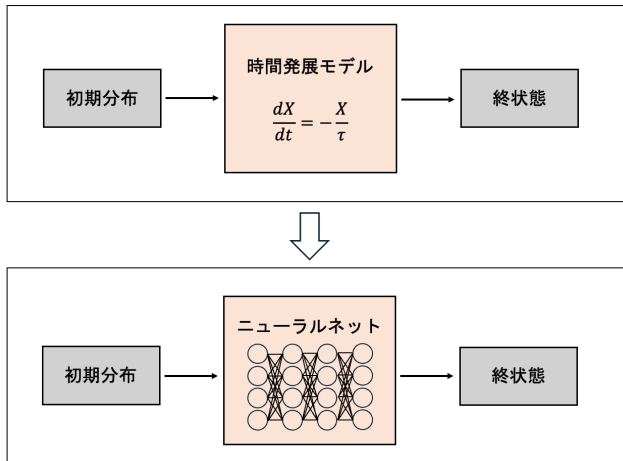
ということがわかった。

また、時間発展モデルを使うと逆問題的に初期条件や進化条件を推定できて、次のことが言える。

- 観測を再現するためには惑星大気はH/He主体でなければならない。
- また初期状態で $X \sim O(1)\%$ 程度の気体を持っている必要がある。
  - → **惑星の大気獲得は原始惑星系円盤から行われなければならない。**
- 観測を再現するためには惑星のコア組成はice/iron-silicateでないといけない。
  - → **アイスラインの外からのimmigrationモデルは否定される。**
- 蒸発は最初の100Myrで起きるとして説明できる。

# 物理モデルを必要としない逆問題の解法： NN と定性的正則化を用いた方法

解析モデルを構築できていない状態で逆問題を解くことはできないか？



# 物理モデルを必要としない逆問題の解法

初期惑星分布を  $(P, R)$  空間上の二次元正規分布と仮定し、パラメタ  $(\mu, \Sigma)$  によって決定づける。

初期惑星分布  $(P_i, R_i) \sim \mathcal{N}(\mu, \Sigma)$  を生成し、ニューラルネットにより進化させる一連のプロセスを「機械学習モデル」と考え、 $(\mu, \Sigma)$  とニューラルネットワークの重み  $W$  を同時に学習する。  
(次ページ参照。)

※ よって、機械学習モデルとしては、惑星数  $n$  を受け入れ、「初期分布生成 → ニューラルネットワークによる進化」を行い、最終惑星分布を出力する。

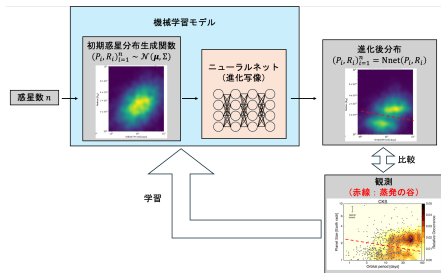
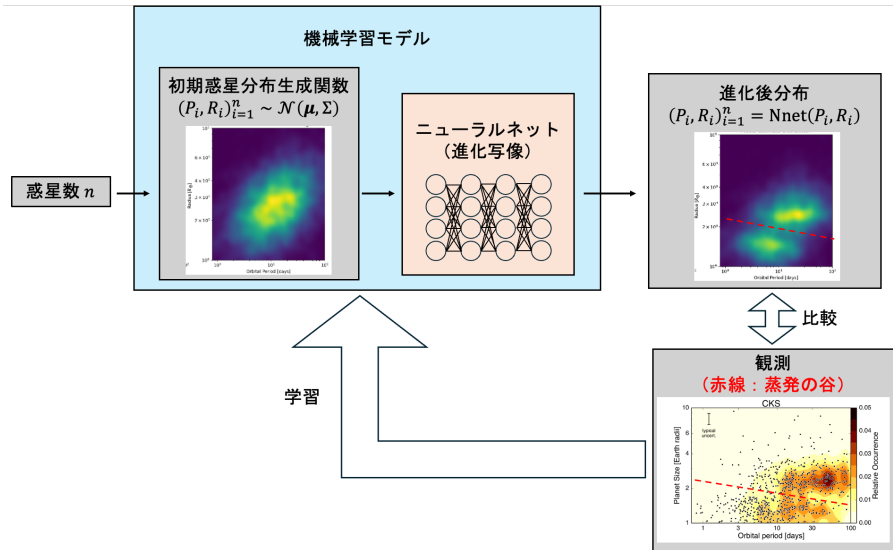


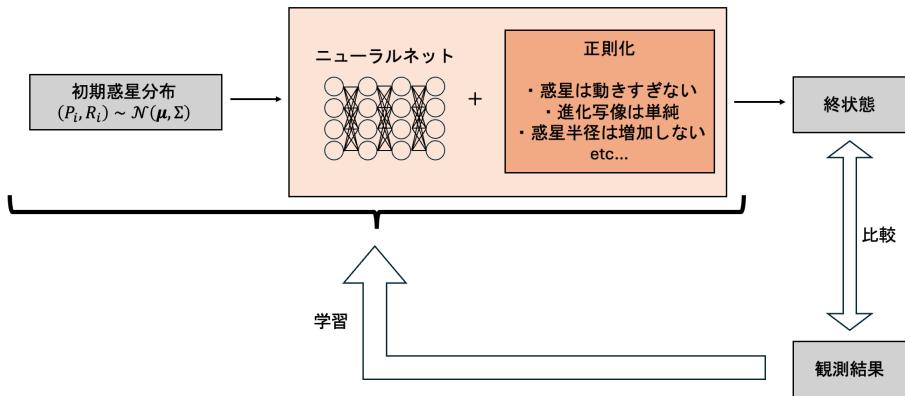
Figure: ※ 次ページに拡大図あり

# 物理モデルを必要としない逆問題の解法

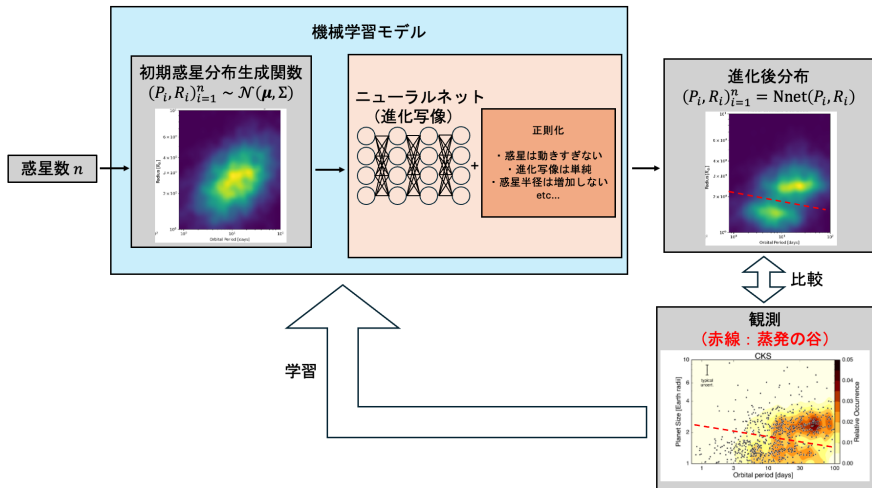


# 正則化条件

- ・初期惑星分布とニューラルネットをまとめて学習する。
- ・ニューラルネットは柔軟すぎて初期惑星分布の動きを吸収してしまう。
  - → 物理モデルに関する直感的な情報を正則化として入れる。



# 正則化条件



# 正則化条件

## 今回導入する正則化条件

- 進化写像は全惑星で共通であり、他の惑星の情報に依存しない

$$P'_i = P_i + f^P(P_i, R_i)$$

$$R'_i = R_i + f^R(P_i, R_i)$$

※ すなわちニューラルネットワークの入出力次元数は 2

- 惑星の半径は減少しかしない

$$f^R < 0$$

※ 学習の上では softplus 関数として実装する。

- 進化関数が複雑すぎない（3次や4次関数のような複数ピークを持つ写像ではない）

$$\int \left| \frac{\partial^n f^x}{\partial (y)^n} \right|^2 dy : \text{small} \rightsquigarrow \sum_i \sum_j \left| \frac{\partial^n f_i^x}{\partial (y_j)^n} \right|^2 : \text{small}$$

※  $x, y = R \text{ or } P$ ,  $n = 3$ , ただし周期については周期方向に二次関数  $\partial^2 f^P / \partial P^2$  について制約する。（惑星の周期変化は特に周期方向に単調であって欲しいので。）

- 惑星は激しく移動・半径変化しない（各惑星の  $R, P$  変化は小さい）

$$\sum_i \left| f_i^P + f_i^R \right|^2 : \text{small}$$

# 損失関数の設計

進化後の惑星の  $(P, R)$  分布と観測で得られている惑星の  $(P, R)$  分布が『近い』ほど小さくなるような損失が望ましい。

一般の多次元離散分布  $\{\mathbf{x}_i\}_{i=1}^N, \{\mathbf{y}_i\}_{i=1}^M$  の「近さ」の指標として次の MMD(Maximum Mean Discrepancy) を用いることができる。

$$\mathcal{L}_{\text{MMD}}(\mathbf{x}, \mathbf{y}) = \frac{1}{N^2} \sum_i^N \sum_j^N K(\mathbf{x}_i, \mathbf{x}_j) + \frac{1}{M^2} \sum_i^M \sum_j^M K(\mathbf{y}_i, \mathbf{y}_j) - \frac{2}{NM} \sum_i^N \sum_j^M K(\mathbf{x}_i, \mathbf{y}_j)$$

ただし、サンプル上の二つの点の近さに対応した量であるカーネル  $K(\mathbf{x}, \mathbf{y})$  の取り方はさまざま考えられるが、ここでは Gaussian RBF Kernel を採用する。

$$K(\mathbf{x}_i, \mathbf{y}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{y}_j\|^2}{2\sigma^2}\right)$$

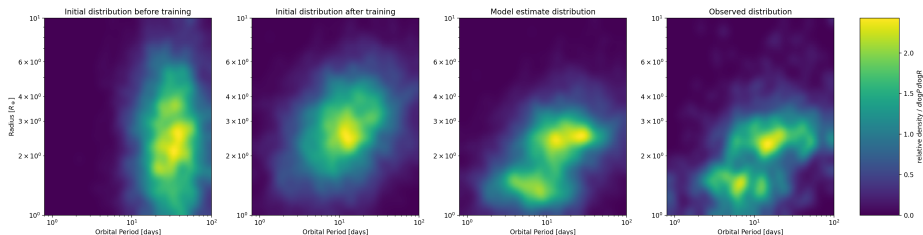
よってニューラルネットで進化させた後の惑星の  $(P, R)$  分布を  $\mathbf{z}_i^{\text{model}} = (P_i, R_i)$ 、観測で得られている惑星の  $(P, R)$  分布を  $\mathbf{z}_j^{\text{obs}} = (P_j, R_j)$  とし、正則化項をまとめて  $\mathcal{R}_{\text{es}}$ 、ハイパーパラメタを  $\lambda$  で表す時、損失関数は次のようになる。

$$\mathcal{L} = \mathcal{L}_{\text{MMD}}(\mathbf{z}^{\text{model}}, \mathbf{z}^{\text{obs}}) + \lambda \cdot \mathcal{R}_{\text{es}}$$

# ハイパーパラメタの選択

ハイパーパラメタ  $\lambda$  は正則化の強さを決めるパラメタである。

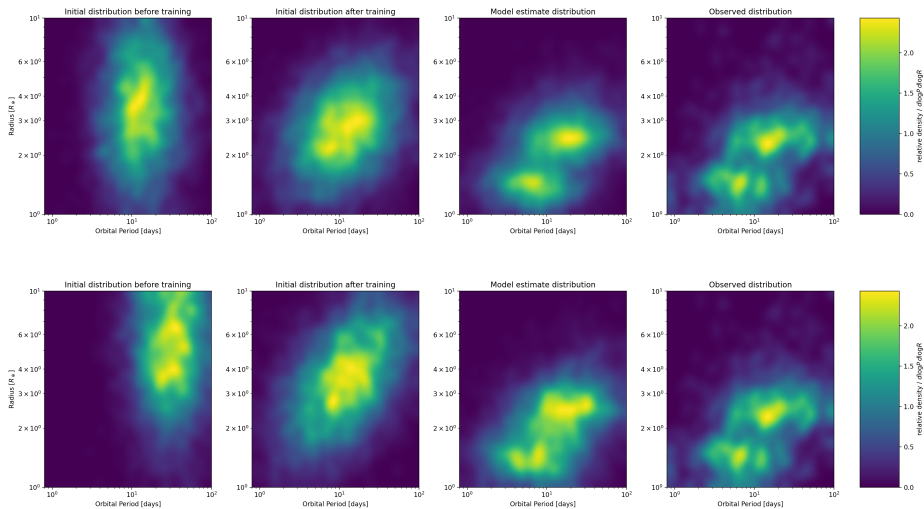
$\lambda$  が小さすぎると、惑星が自由に動いてしまうので初期分布の情報が全てニューラルネットに吸収されてしまう。よって半径の谷はよく再現できる（画像右から二番目）が、初期惑星分布の初期分布（つまり学習前の  $\mu, \Sigma$  の値であり学習の初期条件、画像一番左）をさまざまに変えた時に最終的な初期惑星分布（画像左から二番目）が一意に定まらない。



**Figure:** 左から：学習前の初期惑星分布（学習初期条件）、学習後の初期惑星分布、学習後の初期惑星分布を学習後の NN で進化させて得られる惑星分布、観測で得られている初期惑星分布。

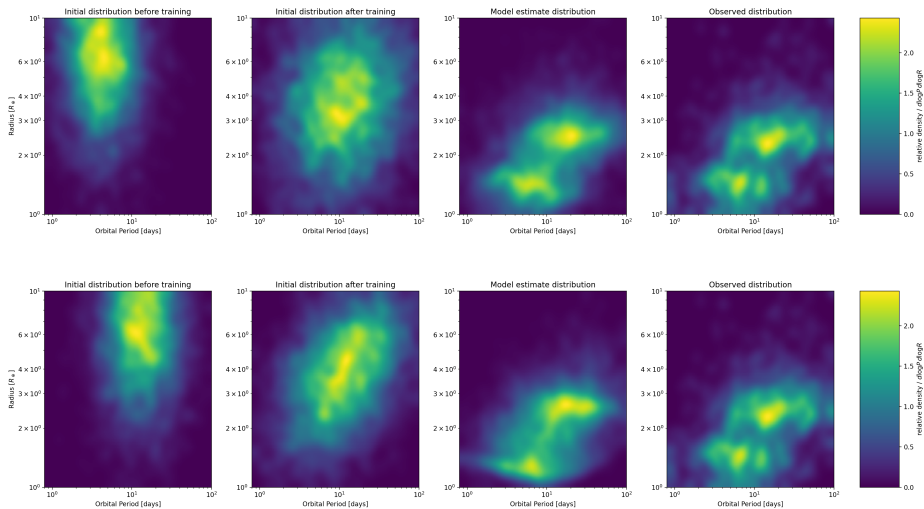
# ハイパーパラメタの選択

$\lambda$  が小さく正則化が弱い場合 :



# ハイパーパラメタの選択

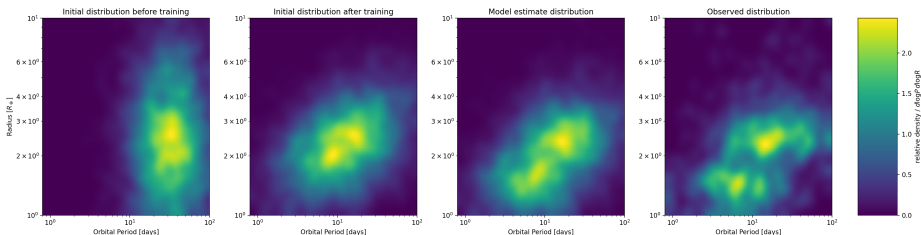
$\lambda$  が小さく正則化が弱い場合：



# ハイパーパラメタの選択

$\lambda$ が大きすぎると、惑星が動きづらくなるので初期惑星分布の初期分布（つまり学習前の  $\mu, \Sigma$  の値であり学習の初期条件、画像一番左）をさまざまに変えた時の最終的な初期惑星分布（画像左から二番目）が一意に定まりやすくなるが、観測の再現度が悪くなり半径の谷の再現が難しくなる（画像右から二番目）。

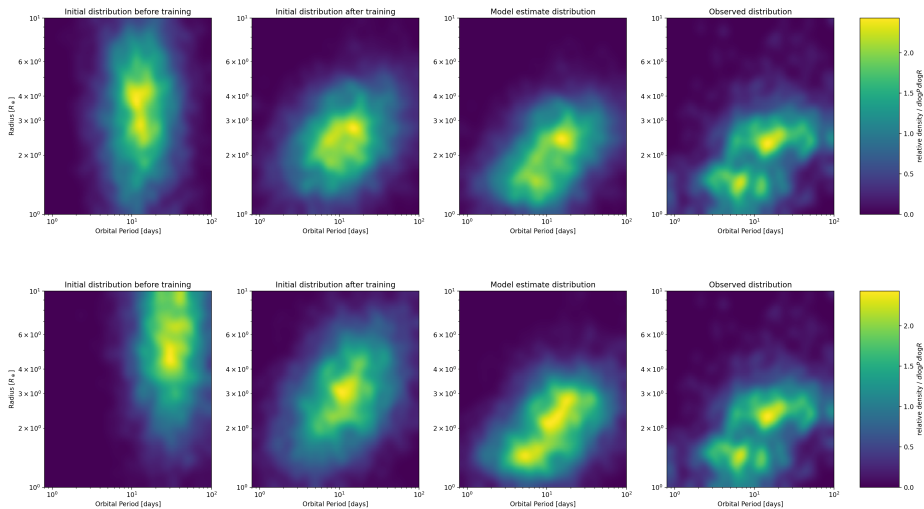
※ 半径の谷の再現度が悪くなるのは割と明白だが、最終的な初期惑星分布が一意に定まりやすくなっているかどうかは諸説ある。（自分は  $\lambda$  を大きくすると最終的な初期惑星分布が割と似てくるように感じているが、あくまで主観的なものである。）



**Figure:** 左から：学習前の初期惑星分布（学習初期条件）、学習後の初期惑星分布、学習後の初期惑星分布を学習後の NN で進化させて得られる惑星分布、観測で得られている初期惑星分布。

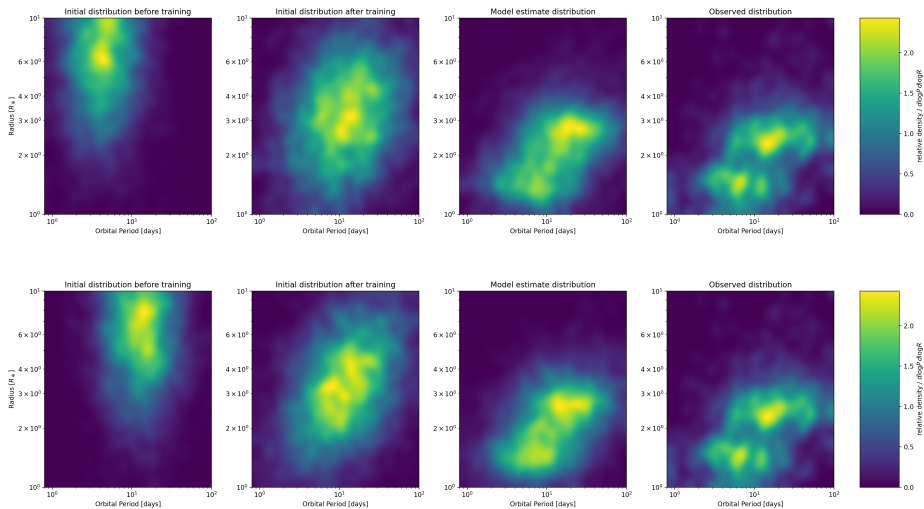
# ハイパーパラメタの選択

$\lambda$  が大きく正則化が強い場合 :



# ハイパーパラメタの選択

$\lambda$  が大きく正則化が強い場合 :



$\lambda$  が極端だと「観測を再現できなくなる」「最終的な初期惑星分布が学習の初期条件に依存してしまい一意に定まりづらくなる」といった現象が見られる。

よって「観測を再現しつつ」「最終的な初期惑星分布が学習の初期条件に依存せず一意に定まる」ような  $\lambda$  を選択する必要がある。

→ 学習初期条件を  $\alpha_1, \alpha_2, \dots, \alpha_K$ 、最終的な初期惑星分布を  $z^0$  で表す時、例えば次のような情報量基準 (Information Criterion) を採用すればいい。

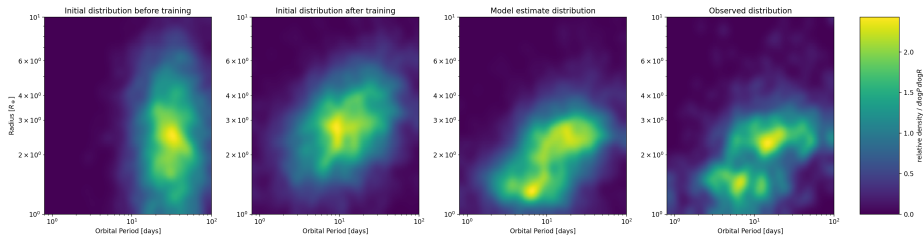
$$\begin{aligned} \text{IC}(\lambda) = & \frac{1}{K} \sum_{1 \leq i \leq K} \mathcal{L}_{\text{MMD}}(z^{\text{model}}, z^{\text{obs}}; \lambda, \alpha_i) \\ & + \beta \times \frac{2}{K(K-1)} \sum_{1 \leq i \leq K} \sum_{1 \leq j \leq K} \mathcal{L}_{\text{MMD}}(z^0(\lambda, \alpha_i), z^0(\lambda, \alpha_j)) \end{aligned}$$

※ 観測を再現するほど第一項が小さくなり、最終的な初期惑星分布が学習初期条件に依存しなくなるほど第二項が小さくなるよう情報量基準を設計した。ただしこれはアドホックに定義を考えたものであり、数学的な厳密性は保証されない。

※  $\beta$  の恣意性は最後まで残る。今回は、第一項と第二項のオーダーがおおよそ一致するように決めた。

# 最終的な解

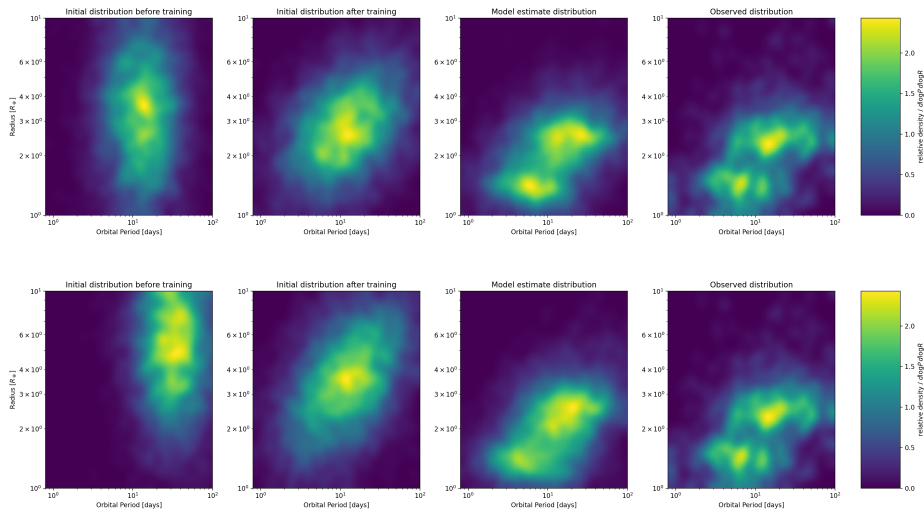
もっとも  $IC(\lambda)$  が小さくなる  $\lambda$  に対し学習の結果の一部を示すと次のよう。



**Figure:** 左から：学習前の初期惑星分布（学習初期条件）、学習後の初期惑星分布、学習後の初期惑星分布を学習後の NN で進化させて得られる惑星分布、観測で得られている初期惑星分布。

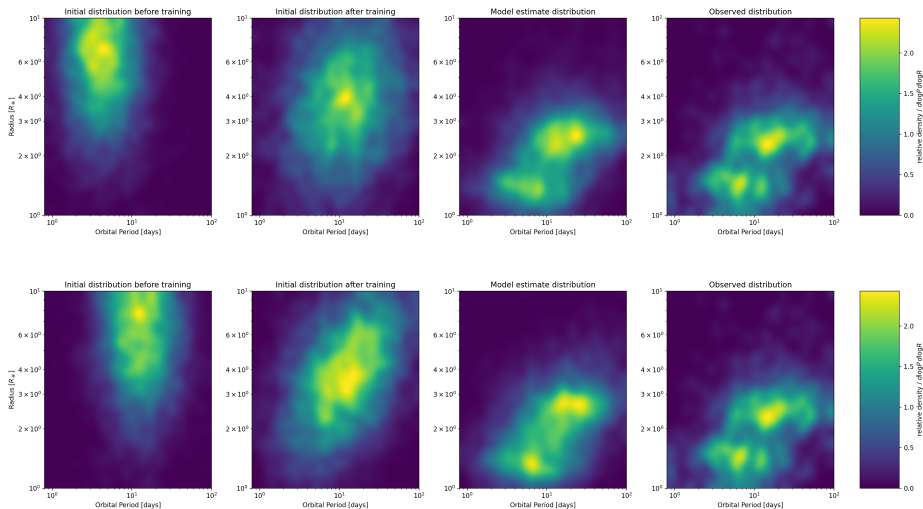
# 最終的な解

情報量基準  $IC(\lambda)$  が最も小さくなる  $\lambda$  :



# 最終的な解

情報量基準  $IC(\lambda)$  が最も小さくなる  $\lambda$  :



# 結果の考察

今回の逆問題の解（惑星の初期 ( $P, R$ ) 分布、画像左から二番目）について、次のことがわかった。

- 今回の方法で得られた惑星の初期 ( $P, R$ ) 分布は、 $\mu^P \sim 10 \text{ days}$ 、 $2R_{\oplus} \sim \mu^R \sim 4R_{\oplus}$  に中心を持つような、 $P, R$  に正の相関を持った二次元正規分布である。
- $R$  についての推定は不確かさが大きい ( $\mu^R$  のばらつきが大きい。)
- $P$  についての推定は不確かさが小さく思える。(正則化の強さ、学習初期条件のばらつきによらず  $\mu^P \sim 10 \text{ days}$  となっている。)
- 初期惑星分布において  $P, R$  の正の相関は、正則化の強さ・学習初期条件のばらつきによらず普遍的に見られる。

総論としては、

- どんなに正則化を強くしても、逆問題の解が学習初期条件の影響を受けてバラついてしまう問題を完全に無くすることができず、手法としての有効性は低いと思われる。
- ただし、正則化の強さ・学習初期条件のばらつきによらず推定される惑星初期分布の周期方向の中心が 10days に強く収束したのは特筆すべきだと思われる。
- 同様に、正則化の強さ・学習初期条件のばらつきによらず推定される惑星初期分布に  $P, R$  について正の相関が見られたことは特筆すべきだと思われる。