arXiv:2602.06492v1 [astro-ph.IM] 6 Feb 2026

作为星系模拟与巡天之间桥梁的神经网络演示

E. Elson[1]

[1]南非贝尔维尔7535，罗伯特·索布克韦路，西开普大学物理与天文学系 2026年2月9日

摘要

本文展示，GAMA（星系与质量组装）巡天中星系的恒星质量，原本通过恒星族群合成建模推导得出，如今仅凭其绝对星等和颜色指数即可准确预测。本研究的一项核心贡献是证明，这个长期存在的推断问题可以通过一种极为简单的机器学习模型解决：仅含一个隐藏层的全连接前馈人工神经网络。该网络仅使用由 Shark 半解析模型生成的合成星系进行训练，并且能够有效迁移至真实观测。在近 3.5 dex 的恒星质量范围内，预测值与 GAMA 光谱能量分布推导质量高度吻合，典型散布为 ~ 0.131 dex。这些结果表明，稳健的恒星质量估算并不需要复杂的深度学习架构，而且在模拟数据上训练的轻量级机器学习模型能够捕捉宽波段测光中编码的主要物理信息。该方法还应用于 17 006 个缺少基于光谱能量分布推导质量的 GAMA 星系，并将测光不确定性通过网络传播，从而为推断出的恒星质量提供相应的误差估计。总而言之，本研究为星系演化研究中的模拟到观测迁移学习，建立了一条计算高效且概念清晰的途径。

1 引言

估算星系的恒星质量仍是河外天文学的一项核心挑战。恒星质量并非直接可观测量；相反，通常需要通过光谱能量分布（SED）拟合来推断，而这需要对恒星形成历史、金属丰度、尘埃消光和恒星初始质量函数等因素作出假设 [Pforr等人，2012，Conroy，2013]。不同研究采用的这些建模选择可能差异很大，从而引入系统不确定性，并使恒星质量估计本质上依赖于模型[Courteau等人，2014]。

相比之下，宇宙学星系形成模拟将恒星质量作为一项基本结果进行预测，其模型经过校准，能够再现观测到的恒星质量函数（例如，Schaye等人 2015，Pillepich等人 2018，Davé等人 2019）。此类模拟基于物理原理，对星系增长、颜色和光谱性质作出预测，并自然地编码了光度与恒星质量之间的关系。因此，现代星系演化研究中一个新兴的问题是，如何将模拟中的理论信息高效地迁移到观测领域，以帮助或加速恒星质量估算。难点在于弥合模拟星系与真实星系之间的差距，因为二者未必分布在可观测参数空间的相同区域。

机器学习为跨越这一鸿沟提供了一条途径。具体而言，前馈人工神经网络（人工神经网络）能够学习测光之间复杂的非线性映射

输入与星系的物理性质。在 Elson [2025], 中表明，即便是一个极其简单的网络——由具有单个隐藏层的全连接前馈人工神经网络构成，且仅以合成星系作为训练对象，这些合成星系生成自SHARK 半解析模型 [Lagos 等人， 2018], ——仍能够仅利用宽带测光和颜色，以高保真度恢复模拟恒星质量。该项工作不仅表明，模拟星系包含足以预测恒星质量的可观测信息，还表明此项任务并不需要复杂架构（例如深度卷积神经网络或编码器—解码器框架）。至关重要的是，研究提出此类模型或许可以迁移到真实数据。

因此，本研究的重点不是开发新模型，而是测试此前由 Elson [2025] 训练的人工神经网络能否直接应用于真实观测数据。除针对观测覆盖范围差异对输入特征进行少量调整外，网络架构和训练集（模拟星系属性）均保持不变。核心目标是评估基于模拟训练的网络能否恢复星系与质量组装（GAMA）巡天中真实星系的恒星质量 [Driver 等人， 2011], 并通过与基于光谱能量分布推导的质量进行比较，以及将其应用于缺少基于光谱能量分布估计值的星系，评估其性能。由此，本研究直接检验了用于恒星质量推断的、从模拟到观测的迁移学习。

2 GAMA 数据与人工神经网络

GAMA星系的恒星质量取自 `StellarMassesLambdarv24` 表格，可通过 GAMA DR4 架构浏览器获取[1]。这些质量是通过宽波段光谱能量分布拟合得出的 [Taylor 等人， 2011], 所用测光数据包括光学 SDSS （ DR7 ）和近红外英国红外深空巡天大面积巡天（ DR4 ）。观测到的光谱能量分布与 Bruzual 和 Charlot [2003], 的恒星族群模型进行比较，这些模型涵盖一系列恒星形成历史、金属丰度和尘埃消光，并假设采用 Chabrier [2003] 初始质量函数。对恒星质光比进行贝叶斯拟合，得到的恒星质量的典型不确定度为 dex。

在Elson [2025], 中展示了，SHARK 半解析模型生成的星系从远紫外到远红外波段的绝对星等，可作为输入特征输入简单的前馈神经网络，以预测恒星质量。本研究将同一训练好的模型应用于观测数据。具体而言，将测得的 GAMA星系绝对星等和颜色指数作为网络的输入，以恢复通过光谱能量分布拟合得出的恒星质量。

理想情况下，Elson [2025] 中使用的同一组特征也适用于 GAMA星系，这样便可直接应用该网络。但实际应用中存在两个限制。首先，GAMAgkvScienceCatv02 表未提供 Spitzer $W3$ 和 $W4$ 星等，因此$W2-W3$ 也无法作为可用的颜色指数。其次，为了使网络可靠运行，每个 GAMA 特征的分布都必须处于相应SHARK 训练数据所涵盖的范围内。除 $FUV-NUV$和 $W1-W2$之外，所有可用的 GAMA 特征均满足这一条件；这两项远超训练范围，因此被排除。对于少数其他特征，只有少量 GAMA星系略微超出了SHARK 的范围。在这些情况下，我们剔除了受影响的星系，使输入集仍能充分代表模拟特征空间。

将 GAMA星系与 `gkvScienceCatv02` 中的测光数据以及 `StellarMassesLambdarv24`中的恒星质量进行交叉匹配后，最终获得了一个包含 71 171 个星系的样本。仅保留满足 $NQ > 2$ （归一化红移质量，确保光谱红移可靠）且$SC > 3$ （科学样本类别，限定为主 $r < 20.5$ 巡天选择中位于天空可靠区域的唯一星系）的星系。星等和颜色的分布如图 1所示。利用光度距离将视星等转换为绝对星等：

[1] https://www.gama-survey.org/dr4/schema/table.php?id=691

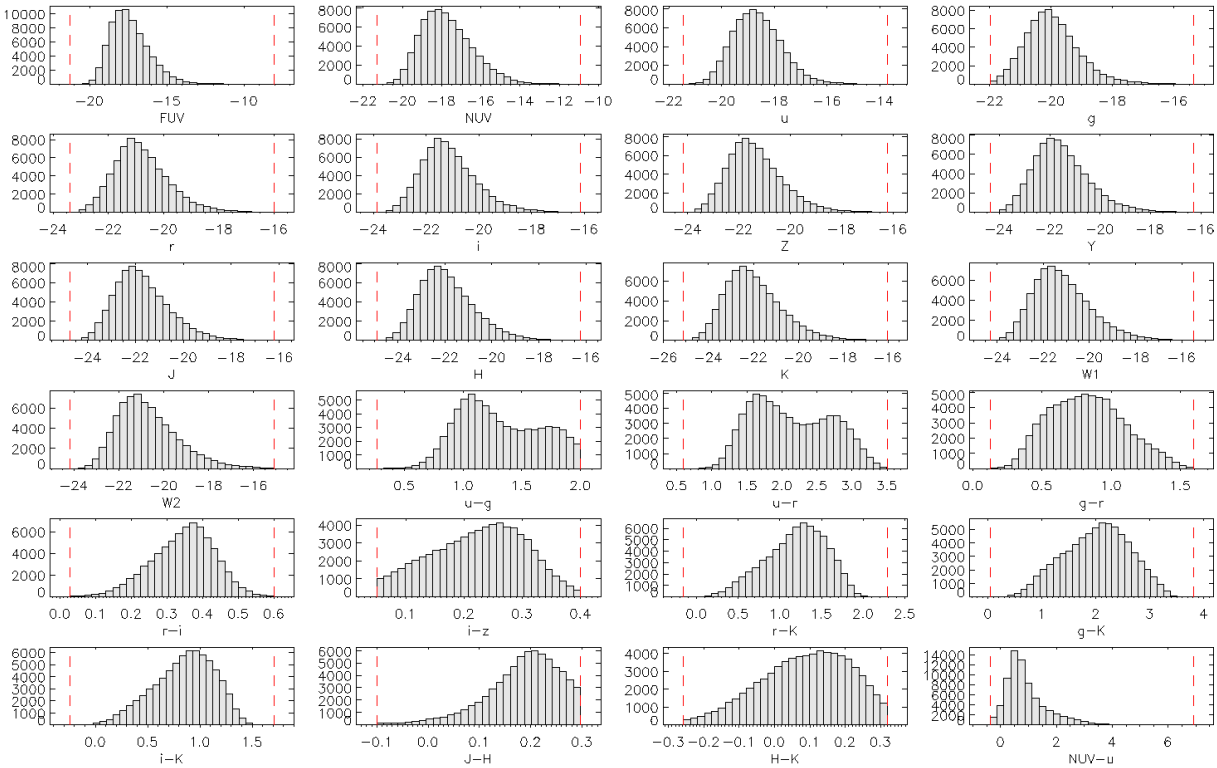


图1：本研究中用作 GAMA 星系输入特征的 24 个宽波段绝对星等和颜色指数（均采用 AB 星等系统）的分布。每个面板中的红色虚线表示 Shark 训练样本中相应特征的最小值和最大值。为确保所有输入都落在人工神经网络的训练范围内，排除了超出这些范围的 GAMA 星系。

光度距离，$D_{\rm L} = (1+z)D_{\rm comov}$，其中 z 是光谱红移，$D_{\rm comov}$是共动距离。

由于 $W3$、$W4$、$W2-W3$、$W1-W2$和 $FUV-NUV$特征不可用或不兼容，因此 Elson [2025] 网络使用可用的 24 个星等重新训练[2]和颜色[3]。此外，为了确保覆盖范围延伸至 $\sim 10^{11.5}\ M_\odot$，本研究移除了该研究中用于剔除大质量 AGN 主导系统的核球与总质量比筛选条件（$B/T < 0.65$）。经过这些调整后，重新训练的人工神经网络在 ~ 0.117 Shark 星系上的误差为，而Elson [2025]报告的误差为 0.087 dex。性能有所下降是意料之中的，尤其是因为 $FUV-NUV$已被证明是一种有效的诊断颜色，原始人工神经网络正是通过它隐式地学习了紫外衰减行为。

3 结果

图 2的上图展示了人工神经网络预测的恒星质量与通过宽波段光谱能量分布拟合得到的恒星质量之间的关系。黑色实线表示一比一关系，红色曲线则是对全部 71 171 个独立星系测量值（而非二维分布分箱数据）进行二阶多项式拟合的结果。显然，将仅使用合成 SHARK 星系训练的人工神经网络直接应用于真实 GAMA 数据，效果非常出色。在所探测的整个质量范围内，预测的恒星质量都与基于光谱能量分布得到的数值高度吻合。最多 ~ 0.1 dex 的微小且几乎恒定的偏移，可能是因为

[2]GALEX FUV和 NUV；SDSS u, g, r, i, z；VISTA Y, J, H, Ks；以及 Spitzer $W1, W2$.

[3]$u-g, u-r, g-r, r-i, i-z, r-Ks, g-Ks, i-Ks, J-H, H-Ks, NUV-u$。

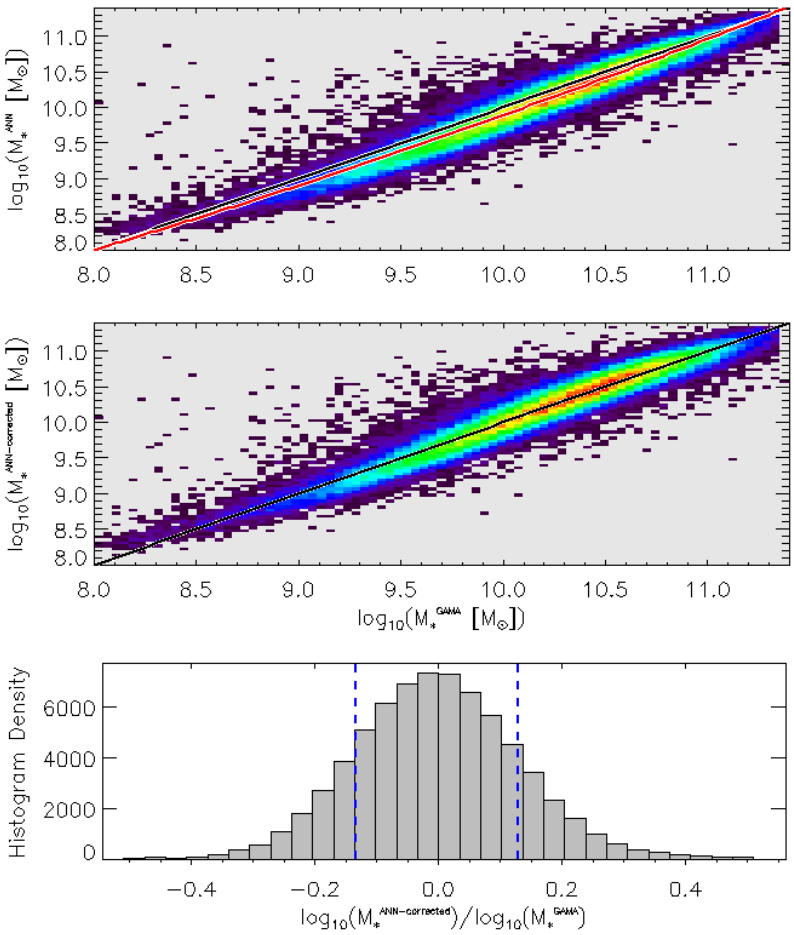


图 2：**上：** 71,171 个 GAMA星系的原始 ANN预测的恒星质量与基于光谱能量分布推导的质量的比较。实线表示一比一关系，红色曲线表示对数据进行的多项式拟合。**中：** 减去拟合的残差趋势后得到的校正 ANN 预测值；此举消除了中位数偏移，使关系与一比一线对齐。**下：** 校正后残差的分布，其中 16–84 百分位范围的一半（蓝色虚线所示）为 0.135 dex。总的来说，这些图表明，ANN 能够非常接近地复现基于SED推导的恒星质量，且仅存在较小、可校正的偏差，整体散布较低。

人工神经网络学习的是 Shark 中编码的质量–光度关系，而 GAMA 恒星质量反映了恒星族群合成建模中所包含的假设。这两个框架不一定会以完全相同的方式将颜色映射为恒星质量。我们使用残差分布的 16–84 百分位范围的一半来量化预测散布（对于近似高斯数据，这是 1σ 散布的稳健估计量），所得数值为 0.135 dex。

顶部面板中黑色的一比一线与红色多项式拟合之间的差值，可以用来量化中位数偏差。该残差在 ~ -0.11 dex 处达到最大值，对应的恒星质量为 $M_* \approx 10^{9.7}\ M_\odot$，略低于 GAMA 的恒星质量中位数 $\approx 10^{10.2}\ M_\odot$。这个平滑的残差函数可用于校正任何 ANN预测的恒星质量。对所有预测值应用此校正后，系统性偏移便会消除。。图 2 的第二个面板展示了校正后的 ANN 质量与基于光谱能量分布推导的质量之间的关系，其中位数重新与一比一关系对齐。最后一个面板展示了校正后的 ANN 预测值与基于光谱能量分布推导的恒星质量之间的残差分布。请注意，正如简单回归偏差所预期的那样，原始比较中的散布仍然保留

校正。

这些结果表明，仅使用模拟星系的测光数据训练的人工神经网络，只需进行少量调整，就能高度准确地复现真实星系的恒星质量。尽管该模型在训练期间从未接触过观测数据，却展现出良好的泛化能力，能够准确匹配基于光谱能量分布推导的恒星质量。这为星系演化研究中未来开展从模拟到观测的迁移学习提供了令人鼓舞的概念验证。

4 应用于缺少基于光谱能量分布推导的恒星质量的 GAMA星系

作为进一步的演示，我们将训练好的人工神经网络应用于缺少基于光谱能量分布推导的恒星质量的 GAMA星系。通过在 `gkvScienceCatv02`目录中筛选满足 SC >4、但未出现在 `StellarMassesLambdarv24` 中的源，我们确定了这 17 006 个天体。对于每个星系，我们提取了人工神经网络所需的输入通量，以及文献中给出的测光不确定性。为将这些误差传播到预测的恒星质量中，我们将每个通量f扰动其对应不确定性的一半，生成两个实现 $f+\delta f/2$ 和 $f-\delta f/2$，其中 δf 为报告的通量误差。我们对每个星系使用 $\{f-\delta f/2,\ f,\ f+\delta f/2\}$ 评估人工神经网络三次，从而得到中心质量估计值及其相应的上下限。

该样本预测的恒星质量如图 3上图所示。将其与 WISE $W1$ 星等进行比较后，可以看到在对数–对数空间中呈现出明显且高度线性的关系，这与预期相符，因为 $W1$ 已知 [Jarrett 等人指出，2023]通量能够反映恒星质量。即使对于没有基于光谱能量分布得出的质量值的星系，这一表现也让我们有信心认为，人工神经网络给出的恒星质量估计在物理上是合理的。

图 3 下图展示了 $\log_{10}(M_*)$ 的传播不确定性随预测质量的变化情况。不确定性范围为 ~ -0.05 至 $+0.05$ dex，并在质量范围的中间附近呈现典型的峰值。这一表现符合物理预期：中等质量星系的恒星形成历史、尘埃含量和淬火状态差异很大，因此其测光质量光度比比低质量恒星形成系统或高质量宁静系统更加退化。

通量传播不确定性显著小于此前对已知光谱能量分布质量的星系测得的~ 0.131 dex人工神经网络内禀散布。由于内禀散布已经包含观测噪声、建模不确定性以及从模拟到观测的迁移局限，因此不应将通量驱动的不确定性与其平方和开方。相反，~ 0.05 dex的扰动范围仅反映预测对通量测量误差的敏感性，而 ~ 0.131dex的散布仍是占主导地位的不确定性分量。因此，ANN预测的恒星质量具有一个非常保守的不确定性估计，即 $0.131+0.05\sim 0.18$ dex。

作为背景参考，将基于人工神经网络的恒星质量不确定性与采用更传统方法得到的不确定性进行比较很有帮助。通过光谱能量分布建模得出的恒星质量，通常存在 ~ 0.2–0.3 dex 的不确定性，这是由于恒星年龄、金属丰度、尘埃消光和恒星形成历史之间存在简并 [例如，Walcher 等人，2011年，Conroy， 2013]。此外，不同的基于光谱能量分布的质量估计之间的比较往往会发现不可忽略的系统性偏移 [例如，Elson， 2024],这强调了此类质量本身也取决于模型。

经验标度关系提供了其他估算方法，但同样受到本征散布的限制。例如，重子塔利–费舍尔关系的估算不确定度为 ~ 0.1–0.4 dex[Bradford 等人，2016年，Lelli 等人， 2016]。在此背景下，本文提出的基于人工神经网络的方法仅需宽带测光，便能达到与常用方法相当的性能。即使采用保守的不确定度估计 ~ 0.18 dex，结果也表明，经模拟训练的神经网络可作为一种实用且互补的工具，用于大型星系巡天中的恒星质量估算。

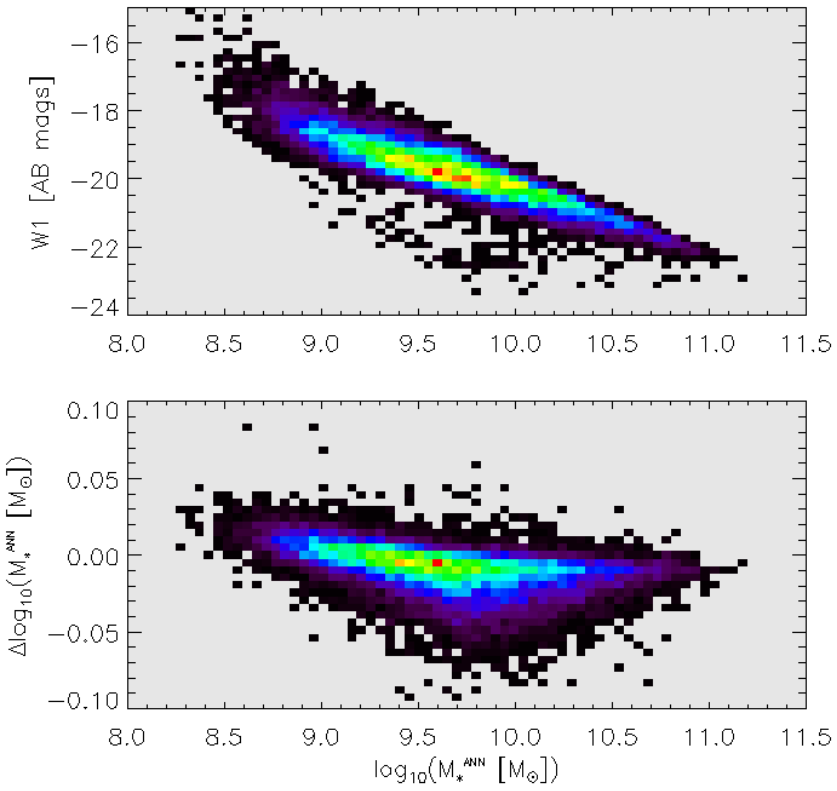


图3：**上图：**没有光谱能量分布估计值的 GAMA星系，其 WISE $W1$ 星等与人工神经网络预测的恒星质量之间的关系，显示出清晰的线性关系，与 $W1$ 追踪恒星质量的结果一致。**下图：** $\log_{10}(M_\star)$的传播不确定性随预测质量的变化，通常为± 0.05 dex，在中等质量处最大；此处星系族群在尘埃含量和恒星形成历史方面最多样。

5 总结

本研究表明，仅使用模拟数据训练的极其简单的机器学习模型，也能够准确推断星系恒星质量。具体而言，研究表明，仅包含一个隐藏层的全连接前馈人工神经网络——完全基于 Shark 半解析模型生成的合成星系进行训练——仅使用宽带星等和颜色指数作为输入特征，即可对 GAMA 巡天中的真实星系进行稳健的恒星质量估算。据作者所知，现有研究尚未证明如此低复杂度的神经网络架构能够在如此宽的动态范围内实现相当的性能。

在进行少量调整以确保模拟数据与观测数据的特征范围相匹配后，该网络能够在约 3.5 dex 的范围内成功重现基于光谱能量分布推导的恒星质量，典型散布为 $\lesssim 0.131$ dex。这一结果直接表明，复杂的神经网络架构并非实现稳健恒星质量推断的必要条件；完成这一任务所需的主要物理信息已经编码在宽带测光数据中。随后，研究将该模型应用于 17 006 个缺乏基于光谱能量分布的质量估计的 GAMA 星系，并将测光不确定性通过网络进行传播，从而得到仅为 ~ 0.05 dex 的典型恒星质量不确定性。总体而言，这些结果建立了一个计算高效且概念透明的从模拟到观测的迁移学习框架，能够快速估算大规模星系样本的恒星质量，并显著降低在星系演化研究中应用机器学习技术的门槛。

参考文献

Jeremy D. Bradford、Marla C. Geha 和 Frank C. van den Bosch。《滑坡：线宽重子塔利-费舍尔关系中的系统不确定性》。*ApJ*，832(1):11，2016年11月。doi: 10.3847/0004-637X/832/1/11。G. Bruzual 和 S. Charlot。以 2003 年的分辨率进行恒星族群合成。《皇家天文学会月报》，344 (4):1000–1028，2003年10月。doi: 10.1046/j.1365-8711.2003.06897.x。Gilles Chabrier。《银河系恒星与亚恒星初始质量函数》。《太平洋天文学会出版物》，115(809):763–795，2003年7月。doi: 10.1086/376392。Charlie Conroy。星系全波段光谱能量分布的建模。《天文学与天体物理学年评》，51(1):393–455，2013年8月。doi: 10.1146/annurev-astro-082812-141017。Stéphane Courteau、Michele Cappellari、Roelof S. de Jong、Aaron A. Dutton、Eric Emsellem、Henk Hoekstra、L. V. E. Koopmans、Gary A. Mamon、Claudia Maraston、Tommaso Treu 和 Lawrence M. Widrow。星系质量。《现代物理评论》，86(1):47–119，2014年1月。doi: 10.1103/RevModPhys.86.47。Romeel Davé、Daniel Anglés-Alcázar、Desika Narayanan、Qi Li、Mika H. Rafieferantsoa 和 Sarah Appleby。SIMBA：包含黑洞增长和反馈的宇宙学模拟。《皇家天文学会月报》，486(2):2827–2849，2019年6月。doi: 10.1093/mnras/stz937。S. P. Driver、D. T. Hill、L. S. Kelvin、A. S. G. Robotham、J. Liske、P. Norberg、I. K. Baldry、S. P. Bamford、A. M. Hopkins、J. Loveday、J. A. Peacock、E. Andrae、J. Bland-Hawthorn、S. Brough、M. J. I. Brown、E. Cameron、J. H. Y. Ching、M. Colless、C. J. Conselice、S. M. Croom、N. J. G. Cross、R. de Propris、S. Dye、M. J. Drinkwater、S. Ellis、Alister W. Graham、M. W. Grootes、M. Gunawardhana、D. H. Jones、E. van Kampen、C. Maraston、R. C. Nichol、H. R. Parkinson、S. Phillipps、K. Pimbblet、C. C. Popescu、M. Prescott、I. G. Roseboom、E. M. Sadler、A. E. Sansom、R. G. Sharp、D. J. B. Smith、E. Taylor、D. Thomas、R. J. Tuffs、D. Wijesinghe、L. Dunne、C. S. Frenk、M. J. Jarvis、B. F. Madore、M. J. Meyer,

M. Seibert、L. Staveley-Smith、W. J. Sutherland 和 S. J. Warren。星系与质量汇编（GAMA）：巡天诊断与核心数据发布。《皇家天文学会月报》，413(2):971–995，2011年5月。doi: 10.1111/j.1365-2966.2010.18188.x。E. Elson。关于 ALFALFA星系的恒星比角动量、质量和有效表面亮度之间存在紧致平面关系。《皇家天文学会月报》，527(1):931–941，2024年1月。doi: 10.1093/mnras/stad3316。E. Elson。利用简单神经网络从宽带星等快速准确地预测恒星质量：在模拟的恒星形成星系中的应用。*RAS* 技术与仪器，4:rzaf029，2025年1月。doi: 10.1093/rasti/rzaf029。T. H. Jarrett、M. E. Cluver、Edward N. Taylor、Sabine Bellstedt、A. S. G. Robotham 和 H. F. M. Yao。新广域红外巡天探测器的恒星质量标定。*ApJ*，946(2):95，2023年4月。doi: 10.3847/1538-4357/acb68f。

Claudia del P. Lagos、Rodrigo J. Tobar、Aaron S. G. Robotham、Danail Obreschkow、Peter D. Mitchell、Chris Power 与 Pascal J. Elahi。Shark：介绍一种开源、免费且灵活的星系形成半解析模型。《皇家天文学会月报》，481(3):3573–3603，2018年12月。doi: 10.1093/mnras/sty2440。

Federico Lelli、Stacy S. McGaugh 与 James M. Schombert。重子塔利-费舍尔关系的小散布。*ApJL*，816(1):L14，2016年1月。doi: 10.3847/2041-8205/816/1/L14。

Janine Pforr、Claudia Maraston 与 Chiara Tonini。通过宽带光谱能量分布拟合恢复星系恒星族性质。《皇家天文学会月报》，422(4):3285–3326，2012年6月。doi: 10.1111/j.1365-2966.2012.20848.x。

Annalisa Pillepich, Volker Springel, Dylan Nelson, Shy Genel, Jill Naiman, Rüdiger Pakmor, LarsHernquist, PaulTorrey, MarkVogelsberger, Rainer Weinberger 和 Federico Marinacci。使用IllustrisTNG模型模拟星系形成。《皇家天文学会月报》，473(3):4077–4106，2018年1月。doi: 10.1093/mnras/stx2656。

Joop Schaye, Robert A. Crain, Richard G. Bower, Michelle Furlong, Matthieu Schaller, Tom Theuns, Claudio Dalla Vecchia, Carlos S. Frenk, I. G. McCarthy, John C. Helly, Adrian Jenkins, Y. M. Rosas-Guevara, Simon D. M. White, Maarten Baes, C. M. Booth, Peter Camps, Julio F. Navarro, Yan Qu, Alireza Rahmati, Till Sawala, Peter A. Thomas 和 James Trayford。EAGLE 项目：模拟星系及其环境的演化与组装。《皇家天文学会月报》，446(1):521–554，2015年1月。doi: 10.1093/mnras/stu2058。

Edward N. Taylor、Andrew M. Hopkins、Ivan K. Baldry、Michael J. I. Brown、Simon P. Driver、Lee S. Kelvin、David T. Hill、Aaron S. G. Robotham、Joss Bland-Hawthorn、D. H. Jones、R. G. Sharp、Daniel Thomas、Jochen Liske、Jon Loveday、Peder Norberg、J. A. Peacock、Steven P. Bamford、Sarah Brough、Matthew Colless、Ewan Cameron、Christopher J. Conselice、Scott M. Croom、C. S. Frenk、Madusha Gunawardhana、Konrad Kuijken、R. C. Nichol、H. R. Parkinson、S. Phillipps、K. A. Pimbblet、C. C. Popescu、Matthew Prescott、W. J. Sutherland、R. J. Tuffs、Eelco van Kampen 和 D. Wijesinghe。《星系与质量汇编》（GAMA）：恒星质量估计。《皇家天文学会月报》，418(3):1587–1620，2011年12月。doi: 10.1111/j.1365-2966.2011.19536. x。

Jakob Walcher、Brent Groves、Tamás Budavári 和 Daniel Dale。拟合星系的积分光谱能量分布。《天体物理学与空间科学》，331(1)：1–51，2011年1月。doi：10.1007/s10509-010-0458-z。