AGN-controlled cooling in elliptical galaxies

P. N. Best,1* C. R. Kaiser,2 T. M. Heckman3 and G. Kauffmann4

1Institute for Astronomy, Royal Observatory Edinburgh, Blackford Hill, Edinburgh EH9 3HJ
2School of Physics & Astronomy, University of Southampton, Southampton SO17 1BJ
3Department of Physics & Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA
4Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748 Garching, Germany

Accepted 2006 February 6. Received 2006 January 13; in original form 2005 December 21

ABSTRACT

A long-standing problem for models of galaxy formation has been the mismatch between the predicted shape of the mass function of dark matter halos and the observed shape of the luminosity function of galaxies. The number of massive haloes is predicted to decrease as a power law (N ∝ M−2) out to very large masses, while the galaxy luminosity function cuts off exponentially at luminosities above L∗. This implies that the efficiency with which gas cools onto massive systems is lower than expected. This letter investigates the role of radio-loud active galactic nuclei (AGN) in continually re-heating the cooling gas. By combining two observational results, the time-averaged energy output associated with recurrent radio source activity is determined, as a function of the black hole mass of the host galaxy: H = 1021.4 (MBH/M⊙)1.6 W. It is shown that for massive elliptical galaxies this radio-source heating balances the radiative energy losses from the hot gas surrounding the galaxy. The recurrent radio-loud AGN activity may therefore provide a self-regulating feedback mechanism capable of controlling the rate of growth of galaxies.

Key words: galaxies: active – galaxies: evolution – galaxies: stellar content – accretion – radio continuum: galaxies.

1 INTRODUCTION

In the widely accepted hierarchical clustering models of galaxy formation, the growth of structure is controlled by the gravitational collapse of haloes of dark matter, which merge together and accrete more mass to form progressively larger structures. These models have provided an excellent framework for describing the distribution of mass in the Universe. The baryonic material, out of which stars and galaxies are formed, is superimposed upon this dark matter distribution. As gas falls into the gravitational potential well of dark matter haloes, it is initially shock-heated to the virial temperature of the halo. This energy is then radiated away, after which the gas cools and condenses, and begins to form stars. The properties of the resultant galaxies can be determined using semi-analytic models (e.g. White & Frenk 1991); these adopt simple prescriptions for ongoing physical processes, such as cooling of the gas, star formation, chemical enrichment, merging of galaxies and feedback mechanisms, in order to predict the properties and distribution of galaxies.

The galaxy luminosity function offers a powerful probe of the process of galaxy formation. In early, simple, semi-analytic models, it was found that the predicted shape of the galaxy luminosity function differed greatly from that which is observed: many fewer galaxies are observed than were predicted at both high and low masses. It was soon discovered that the problem of the lack of faint galaxies could be solved by including in the model the heating effects of supernovae (e.g. White & Frenk 1991) and photoionization by the cosmic background (Efstathiou 1992), which both act to reduce star formation efficiency at the low-mass end. The solution to the deficit of bright galaxies was originally proposed to be due to cooling inefficiencies in massive galaxies, but these effects were subsequently shown to be insufficient. In order to reconcile the semi-analytic models with observations, some additional suppression of the gas cooling in massive galaxies was required. The generation of semi-analytic models of the late 1990s introduced this in an ad-hoc manner (e.g. by artificially switching off cooling in the most massive haloes), and were able to provide a very successful description of galaxy properties in the nearby and high-redshift Universe (e.g. Kauffmann et al. 1999; Somerville & Primack 1999; Cole et al. 2000). In order to properly understand galaxy formation, however, it is clearly important to understand the physical processes involved in suppressing cooling in these galaxies.

Active galactic nuclei (AGN) offer a promising solution to this problem (e.g. Benson et al. 2003; Croton et al. 2005; Bower et al. 2005; Scannapieco, Silk & Bouwens 2005). AGN are predominantly found in the most Massive galaxies (e.g. Kauffmann et al. 2003), and are sufficiently energetic that if their output could be efficiently coupled to the gas it would be more than sufficient to re-heat the gas.
and suppress further cooling. Prescriptions for AGN feedback have recently been introduced into semi-analytic models of galaxy formation, and both Croton et al. (2005) and Bower et al. (2005) show that this feedback is able to solve the issue of the bright end of the luminosity function, whilst simultaneously solving other problems of galaxy formation models such as why the most massive galaxies are so red. These prescriptions are necessarily simple in nature, however, and are not well grounded in observation. Indeed, the form of the AGN feedback adopted is very different in the two prescriptions (despite both providing good solutions), indicating that there is still much work to be done in understanding exactly how and when AGN feedback is important.

The focus of this letter is to determine the AGN feedback due to radio-loud AGN. These are of particular interest because the expanding radio source provides a direct way for the AGN output to be coupled to its environment. Indeed, in clusters of galaxies containing powerful radio sources, X-ray observations have revealed bubbles and cavities in the hot intracluster medium, evacuated by the expanding radio source (e.g. Böhringer et al. 1993; Carilli, Perley & Harris 1994; McNamara et al. 2000; Fabian et al. 2003). The (pV) energy associated with these bubbles (which in turn arises from the mechanical energy output of the radio jet) can be sufficient to balance the cooling losses of some clusters, at least for a short period of time (e.g. Fabian et al. 2003; Bırzan et al. 2004). Radio-loud AGN have relatively short lifetimes (10⁷–10⁹ yr), however, and so the balancing of cooling will not be long-lived. Nonetheless, Best et al. (2005b) recently showed that over 25 per cent of the most massive galaxies show radio-loud AGN activity, which implies that AGN activity must be continually re-triggered.

In this Letter, recent observational results are combined to show that the time-averaged effect of heating by recurrent radio source activity balances the radiative energy losses in the hot envelopes of the ellipticals, across a wide range of black hole masses. Section 2 provides a description of the observational results upon which this work is built. In Section 3 the results are discussed, the observations are compared with the prescriptions used in the semi-analytic models, and conclusions are drawn. Throughout this work, the cosmological parameters are assumed to have values of \( \Omega_m = 0.3 \), \( \Omega_\Lambda = 0.7 \), and \( H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1} \).

## 2 METHODOLOGY

### 2.1 The black-hole mass dependent radio luminosity function

Best et al. (2005a) constructed a sample of 2215 radio-loud AGN with redshifts 0.03 < z < 0.3, from the second data release (DR2) of the Sloan Digital Sky Survey (SDSS), by comparing this with a combination of the National Radio Astronomy Observatory (NRAO) Very Large Array (VLA) Sky Survey (NVSS; Condon et al. 1998) and the Faint Images of the Radio Sky at Twenty centimetres (FIRST) survey (Becker, White & Helfand 1995). Best et al. (2005b) compared these AGN with the radio-quiet galaxies from the parent SDSS catalogue to derive the fraction of galaxies that are radio-loud as a function of black hole mass (\( M_{\text{BH}} \)), as determined from the host galaxy velocity dispersion; e.g. Tremaine et al. (2002) and radio luminosity separately. They found that the fraction of galaxies that are radio-loud depends very strongly on black hole mass, going as \( f_{\text{radio-loud}} \propto M_{\text{BH}}^{\alpha} \). On the other hand, the distribution of radio luminosities was found to be independent of black hole mass. Best et al. were able to parametrize this result using a mass-dependent broken power-law model, such that overall the fraction of sources that are radio-loud AGN brighter than some 1.4-GHz radio luminosity \( L \) is given by

\[
f_{\text{radio-loud}} = f_0 \left( \frac{M_{\text{BH}}}{10^6 M_\odot} \right)^\alpha \left[ \left( \frac{L}{L_\text{\star}} \right)^\beta + \left( \frac{L}{L_\text{\star}} \right)^\gamma \right]^{-1}
\]  

with best-fitting parameters of \( f_0 = (3.5 \pm 0.4) \times 10^{-3} \), \( \alpha = 1.6 \pm 0.1 \), \( \beta = 0.37 \pm 0.03 \), \( \gamma = 1.79 \pm 0.14 \) and \( L_\star = (3.2 \pm 0.5) \times 10^{24} \text{ W Hz}^{-1} \).

In a model in which radio source activity is continually re-triggered, this result can instead be interpreted (probabilistically) as the fraction of its time that a galaxy of given black hole mass spends as a radio source brighter than a given radio luminosity.

### 2.2 Radio luminosity versus mechanical energy output

Monochromatic radio luminosity does not provide a good indicator of the mechanical energy output of a radio source. Bicknell (1995) estimates that the kinetic energy output of a radio jet is typically a factor of 100–1000 higher than the total radio luminosity of a radio source. In order to use the above results to determine the time-averaged heating output of radio-loud AGN, it is therefore necessary to derive a conversion between radio luminosity and mechanical energy output.

Bırzan et al. (2004) studied the cavities and bubbles that are produced in clusters and groups of galaxies due to interactions between radio sources and the surrounding hot gas. They derived the (pV) energy associated with the cavities; combining this with an estimate of the age of the cavities provides an estimate of the mechanical luminosity associated with the radio source. Fig. 1 shows the mechanical luminosities derived by Bırzan et al. (summed over all cavities associated with a given cluster or group) plotted against the monochromatic 1.4-GHz radio luminosity of the associated radio sources. As Bırzan et al. point out, there is no single scaling factor between radio luminosity and mechanical luminosity. Nonetheless, it is possible to obtain a reasonable fit to the data using a simple power-law relation (shown as the solid line in Fig. 1):

\[
\frac{L_{\text{mech}}}{10^{39} \text{ W}} = (3.0 \pm 0.2) \left( \frac{L_{1.4 \text{ GHz}}}{10^{24} \text{ W Hz}^{-1}} \right)^{0.40 \pm 0.13}
\]

where the errors are calculated using bootstrap analysis.

---

**Figure 1.** The mechanical luminosity of radio sources, as estimated by Bırzan et al. (2004) from cavities and bubbles in the X-ray gas, versus the monochromatic 1.4-GHz radio luminosity. The solid line shows an error-weighted best fit to the relation (in linear space). The errors on the mechanical luminosities come from different assumptions in deriving the ages of the bubbles (see Bırzan et al. 2004 for details).
The scatter in Fig. 1 is larger than the error bars, which suggests that there must be some additional dependencies in the conversion from radio to mechanical luminosity. This is not surprising since even for a source of fixed jet kinetic power the radio luminosity changes as the source ages (e.g. Kaiser, Dennett-Thorpe & Alexander 1997). Important, however, is that the offset from the best-fitting line in Fig. 1 does not correlate with other properties of the host galaxy (optical luminosity, velocity dispersion, environment), indicating that for all host galaxies, equation (2) provides a reliable measure of the mean mechanical luminosity associated with a source of given radio luminosity.

Combining equations (1) and (2) then gives the (probabilistic) fraction of time that a galaxy of given black hole mass produces a radio source of given mechanical luminosity, and from this the time-averaged mechanical energy output of a galaxy can be derived as a function of its black hole mass. This evaluates to $H = 10^{21.4}(M_{\text{BH}}/M_\odot)^{1.6}\text{W}$. The black hole mass dependence of this heating rate arises entirely from the variation of the radio-loud fraction with black hole mass, whilst the radio luminosity to mechanical luminosity conversion determines the normalization of the relation.

2.3 Gas cooling rates in elliptical galaxy haloes

The radiated energy losses from the haloes of hot gas surrounding elliptical galaxies can be determined using X-ray observations. O’Sullivan, Forbes & Ponman (2001) have compiled a (not complete, but relatively unbiased) catalogue of 401 early-type galaxies, for which they have tabulated the X-ray luminosities ($L_X$, in erg s$^{-1}$), the optical $B$-band luminosities ($L_B$, in units of the solar luminosity in the $B$-band, $L_B/\odot = 5.2 \times 10^{32}$ erg s$^{-1}$), and the T-type morphology parameter (de Vaucouleurs, de Vaucouleurs & Corwin 1976).

The X-ray luminosities were derived using a common spectral model and distance scale, and converted to a pseudo-bolometric band of 0.088 to 17.25 keV; O’Sullivan et al. show that extending this energy band further makes <10 per cent changes to the bolometric X-ray luminosity.

The X-ray luminosity arises from two separate components of the galaxies: the sum of discrete sources within the galaxy (X-ray binaries, discrete stars, globular clusters, etc.) and the cooling of the hot gas. The luminosity of the discrete sources, being stellar in origin, should scale roughly with the optical luminosity of the galaxy, and O’Sullivan et al. estimate that this scales roughly as $L_X \approx 10^{29.5} L_B$. This can dominate in low-mass galaxies. For higher mass galaxies, the X-ray luminosity is dominated by the emission from the hot gaseous haloes, and O’Sullivan et al. find that $L_X \propto L_B^3$ (see also the review by Mathews & Brighenti 2003).

For the current paper, analysis is restricted to those sources with $T \lesssim 4$, that is, bona-fide elliptical galaxies. A plot of $L_X$ versus $L_B$ for these galaxies is shown in Fig. 2; the dotted line on that plot represents the average contribution to the X-ray luminosities from the discrete sources. The residual effectively measures the cooling energy losses from the haloes of the ellipticals, as a function of host galaxy luminosity.

2.4 Black hole mass versus bulge luminosity relation

In order to compare the average radio-AGN heating rates (as a function of black hole mass) with the average gas cooling rates (as a function of $B$-band luminosity), the black hole masses are converted to galaxy luminosities using the relation of McLure & Dunlop (2002): $\log(M_{\text{BH}}/M_\odot) = -0.50 M_B - 2.96$, where $M_B$ is the absolute $B$-band magnitude of the galaxy bulge. Because the sample is restricted strictly to elliptical galaxies, the galaxy luminosity and the bulge luminosity are identical. The $R$-band magnitude is converted to a $B$-band luminosity using the average $B - R$ rest-frame colour for elliptical galaxies, $B - R \approx 1.2$ (e.g. Blanton et al. 2003), and the absolute magnitude of the Sun $(M_B/\odot = 5.48)$.

3 RESULTS AND DISCUSSION

Fig. 2 shows the $L_X$ versus $L_B$ relation for elliptical galaxies from O’Sullivan et al. (2001). The solid line on this plot shows the time-averaged energy input into a galaxy by its recurrent radio source activity. The agreement between these heating and cooling rates across a wide range of host galaxy luminosities (masses) is remarkable. Averaged over time, the energy supplied to a galaxy by its recurrent radio source activity can balance the radiative energy losses from its hot halo; thus, cooling of the gas is suppressed, and the radio source activity may control the rate of growth of the elliptical galaxy.
For this radio-source feedback model to work, it is necessary that the majority of the mechanical energy supplied by the radio activity is dissipated within the host galaxy halo. However, the most powerful radio sources can have linear sizes very much larger than their host galaxies. Combining the local radio luminosity function for AGN (e.g. Best et al. 2005a) with the conversion from radio to mechanical luminosity (equation 2), the contribution to the global mechanical heating rate from sources of different radio luminosities can be derived: this is shown in Fig. 3. It is reassuring that the heating is dominated by low-luminosity sources, $L_{1.4\,\text{GHz}} \lesssim 10^{24}$ W Hz$^{-1}$; such low-luminosity sources tend to be small, and their mechanical energy will indeed be confined predominantly to size-scales of the host galaxy halo. These low-luminosity sources are also ‘on’ for a significant fraction of the time (up to 30 per cent in the most massive galaxies; see Best et al. 2005b), meaning that the heat supply is reasonably continuous, as opposed to being in the occasional short-lived explosive events associated with the (rarer) more powerful sources.

The volume-averaged mechanical energy heating rate due to all radio sources in the nearby Universe can be determined by integrating Fig. 3; because of the flatness of the relation at low radio luminosities, the total heating budget due to radio sources in the nearby Universe depends upon the radio luminosity down to which the radio luminosity function of AGN extends without change of slope. Adopting $L_{1.4\,\text{GHz}} = 10^{25}$ W Hz$^{-1}$ as the lower limit gives a zero-redshift volume-averaged heating rate of $\dot{H} \approx 4.0 \times 10^{31}$ W Mpc$^{-3}$, but this value would nearly double if the slope of the radio luminosity function for AGN remains unchanged down to $10^{18}$ W Hz$^{-1}$.

Regardless of the precise value, the observed local heating rate due to radio-loud AGN is a factor of 10–20 lower than the local density of radio-mode heating in the Croton et al. (2005) semi-analytic model ($9 \times 10^{32}$ W Mpc$^{-3}$, as calculated from their fig. 3 and equation 10). Part of this difference is that the model cooling rates seem to be over-estimated by a factor of 2–3, due to the schematic nature of the underlying cooling flow model (White, Croton, private communication). The remainder of the difference arises because the level of radio heating in the Croton et al. model is tuned to balance cooling in all massive galaxies, including the special case of brightest cluster galaxies at the centres of extreme cooling flows. The observations presented here, however, suggest only that radio source heating balances cooling radiation losses in ‘typical’ elliptical galaxies; there is no evidence for the additional heating that would be required to control cooling on cluster scales. An alternative heating source is therefore required to solve the cluster cooling flow problem. This could still be associated with radio sources, but only if there is an additional mode of powerful radio source activity associated specifically with brightest cluster galaxies; this is beyond the test of the current observations.

The balance between heating and cooling rates in typical elliptical galaxies is likely to occur through a feedback-controlled episodic behaviour of the AGN activity (cf. Binney & Tabor 1995; Kaiser & Binney 2003): the radio source activity heats the surrounding gas; this then cools over time, until eventually some gas begins to condense down on to the central galaxy; a small proportion of that gas is driven down to the central regions of the galaxy, fuelling a new cycle of AGN activity; the resulting radio source re-heats the gas and cuts-off the cooling once again.

One interesting question is why, in this scenario, the radio-loud AGN fraction should scale as $M_{\text{BH}}^{1.6}$. Best et al. (2005b) found that the occurrence of low-luminosity radio-loud AGN activity was completely independent of that of optical (emission-line) AGN activity, and therefore argued that radio-loud AGN were fuelled by the accretion of hot gas (as opposed to the cold gas which fuels optical AGN through a standard thin accretion disc). Croton et al. (2005) showed that the Bondi–Hoyle accretion rate of hot gas is $\dot{m}_{\text{Bondi}} \approx G \mu m_p k T M_{\text{BH}} / \Lambda$ (their equation 28), where $\mu m_p$ is the mean mass of particles in the gas, and $\Lambda$ is the cooling function. For an isothermal gas, $T \propto \sigma^2$, which from the black hole mass versus velocity dispersion relation gives $T \propto M_{\text{BH}} / \Lambda$, which is relatively independent of temperature (and hence black hole mass) at the temperature of elliptical galaxy haloes, and therefore the Bondi accretion rate scales roughly as $M_{\text{BH}}^5$, remarkably similar to the scaling of the radio-loud fraction: the higher proportion of radio-loud AGN in more massive galaxies may therefore arise naturally from the higher accretion rates. The physics underlying this accretion must be more complicated, since the Bondi accretion rate is for continual infall of gas whilst the mass-scaling of the radio-loud AGN fraction represents a scaling of the fraction of time that a given black hole is active, but it is certainly intriguing that the two scale with mass in the same way, and that this mass scaling matches that required in the Croton et al. model to solve the current problems in the semi-analytic models.

In conclusion, therefore, observations show that the time-averaged mechanical energy output associated with radio-source activity in massive galaxies is found to almost exactly balance the cooling radiation losses from the hot haloes of those galaxies, across a wide range of galaxy masses. This implies that the cooling of the gas may be feedback-controlled by the radio activity. This feedback will control the growth rate of all galaxies in haloes in which a quasi-stationary cooling halo of hot gas has been established; in practice this means all dark matter haloes more massive than $2–3 \times 10^{11} M_\odot$ (e.g. Bower et al. 2005; Croton et al. 2005). The mass-scaling of
the mechanical energy output due to radio source activity matches both that predicted if they are fuelled by Bondi accretion of hot gas from the cooling haloes, and that required for radio-AGN feedback to successfully solve the problem of the over-growth of massive galaxies in semi-analytic models.

ACKNOWLEDGMENTS

PNB would like to thank the Royal Society for generous financial support through its University Research Fellowship scheme. The authors thank Simon White and Darren Croton for illuminating discussions about the cooling and heating rates in the semi-analytic models, Ed Pope for useful discussions about the recurrent radio source activity, and the referee for helpful comments.

REFERENCES

Blanton M. R. et al., 2003, AJ, 125, 2348

LaValle M., Itoh T., Feigelson E., 1992, BAAS, 24, 839

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.