Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 18 July 2002 (MN LATEX style file v1.4)

Consistent β values from density-density and velocity-velocity comparisons

Saleem Zaroubi[1], Enzo Branchini[2], Yehuda Hoffman[3], Luiz N. da Costa[4]
[1] *Max Planck Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85741 Garching, Germany.*
[2] *Dipartimento di Fisica dell'Universitá degli Studi "Roma TRE", Via della Vasca Navale 84, I-00146, Roma, Italy.*
[3] *Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel.*
[4] *European Southern Observatory, Karl-Schwarschild Strasse 2, 85741, Garching, Germany.*

18 July 2002

ABSTRACT

We apply a new algorithm, called the Unbiased Minimal Variance (hereafter UMV) estimator, to reconstruct the cosmic density and peculiar velocity fields in our local universe from the SEcat catalog of peculiar velocities comprising both early (ENEAR) and late type (SFI) galaxies. The reconstructed fields are compared with those predicted from the IRAS PSCz galaxy redshift survey to constrain the value of $\beta = \Omega_m^{0.6}/b$, where Ω_m and b are the mass density and the bias parameters. The comparison of the density and velocity fields is carried out within the same methodological framework, and leads, for the first time, to consistent values of β, yielding $\beta = 0.57^{+0.11}_{-0.13}$ and $\beta = 0.51 \pm 0.06$, respectively.

We find that the distribution of the density and velocity residuals, relative to their respective errors, is consistent with a Gaussian distribution with $\sigma \approx 1$, indicating that the density field predicted from the PSCz is an acceptable fit to that deduced from the peculiar velocities of the SEcat galaxies.

Key words: Cosmology: theory – galaxies: clustering, – large–scale structure, large–scale flows.

1 INTRODUCTION

In the gravitational instability scenario (*e.g.*, Peebles 1980), mass density fluctuations and peculiar velocities evolve in an expanding universe under the effect of gravity. If density fluctuations are small, linear theory is valid and a simple relation exists between peculiar velocities, $\mathbf{v}$, and mass density contrast, δ_m:

$$\nabla \cdot \mathbf{v} = -\Omega_m^{0.6} \delta_m, \quad (1)$$

where Ω_m is the mass density parameter. Equation (1) shows why peculiar motions are so important in cosmology: they provide a direct probe of the mass density distribution in the universe. The mass density fluctuation field, δ_m, can be deduced from the galaxy observed density contrast, δ_g, assuming a relation (bias) between the distribution of galaxies and that of the underlying density. The simplest relation suggested in the literature is that of linear bias, namely $\delta_g = b\delta_m$, where b is the linear bias parameter for a given population of mass tracers. This assumption seems to hold on very large (linear) scales and it is supported by both observational evidence (e.g. Baker *et al.*, 1998 and Seaborne *et al.*, 1999) and numerical studies (*e.g.*, Blanton *et al.*, 1999).

Many authors have used galaxies' peculiar velocities and their redshift space positions to estimate the value of $\beta = \Omega_m^{0.6}/b$, under the hypotheses of linear theory and linear biasing. These analyses have been typically carried out using two alternative strategies. In the so-called density-density comparisons a 3-D velocity field and a self-consistent mass density field are derived from observed radial velocities and compared to the galaxy density field measured from large redshift surveys. The typical example is the comparison of the mass density field reconstructed with the POTENT method (Bertschinger & Dekel 1989, Dekel *et al.*, 1990) from the MARK III catalog of galaxy peculiar velocities (Willick *et al.*, 1997a) with the galaxy density field obtained from the IRAS 1.2 Jy redshift catalog (Sigad *et al.*, 1998). The various applications of density-density comparisons to a number of datasets have persistently led to large estimates of β, consistent with unity (see

Sigad *et al.*, 1998 and references therein). The alternative approach is constituted by the velocity-velocity analyses. In this second case the observed galaxy distribution is used to infer a mass density field from which peculiar velocities are obtained and compared to the observed ones. The velocity-velocity methods have been applied to most of the velocity catalogs presently available yielding systematically lower values of β, in the range $0.4 - 0.6$ (see Zaroubi 2002a, for a summary of the most recent results).

Both density-density and velocity-velocity methods have been carefully tested using mock catalogs extracted from N-body simulations. They were shown to provide an unbiased estimate of the β parameter. Yet, when applied to the same datasets, the discrepancy in the β estimates turned out to be significantly larger than the expected errors. Accounting for mildly nonlinear motions (e.g. Sigad *et al.*, 1998 and Willick *et al.*, 1996) or allowing for possible deviations from a pure linear biasing relation consistent with the observational constraints (see discussion in Somerville *et al.*, 2001, Branchini *et al.*2001) does not explain this discrepancy (Berlind, Narayanan and Weinberg 2001). Velocity-velocity comparisons are generally regarded as more reliable as they require manipulation of the denser and more homogeneous, redshift catalog data. Whereas, the density-density comparisons involve manipulation of the noisier and sparser velocity data. In any case both classes of methods are quite complicated and it is hard to understand how systematic errors can arise and propagate through the analysis. Therefore, it is likely that these systematics affect the β parameter estimation.

The purpose of this work is to address, and possibly solve, the density-density *vs.* velocity-velocity dichotomy. We achieve this goal by using the novel Unbiased Minimal Variance estimator, recently proposed by Zaroubi (2002b). The UMV estimator allows one to reconstruct an unbiased cosmological field at any point in space from sparse, noisy and incomplete data and to map it into a dynamically-related cosmic field (*e.g.* to go from peculiar velocities to overdensities). The UMV is applied here to the SEcat catalog of peculiar velocities (Zaroubi 2002b) to reconstruct both the mass density and peculiar velocity fields. These fields are then compared with the analogous quantities predicted from the distribution of IRAS PSCz galaxies (Saunders *et al.*, 2000) of density-density and a velocity-velocity analyses. In Section 2 we briefly review the basics of the UMV estimator. The SEcat and PSCz catalogs are presented in Section 3. Error estimation from mock catalogs is described in § 4. The density and velocity fields obtained by applying UMV to SEcat are compared in Section 5 with the analogous quantities deduced from the PSCz dataset. Finally, in Sections 6 we discuss the results and present our conclusions.

2 THE UMV METHOD

The derivation and general properties of the UMV estimator have been already presented and discussed by Zaroubi (2002b). Therefore, we only review its main properties. As in the case of Wiener filter (Wiener 1949, Zaroubi *et al.*, 1995), the UMV introduces a general framework of linear estimation and prediction by minimizing the variance of the estimator with respect to the underlying signal, subject to linear constraints on the data. The solution of the minimization problem yields the UMV estimator, which was shown to be a very effective tool for reconstructing the large scale structure of the universe from incomplete, noisy and sparse data (Zaroubi 2002b) .

One of the main drawbacks of the Wiener filter is that it suppresses the amplitude of the estimated signal. The suppression factor is roughly equal to the $Signal^2/(Signal^2 + Noise^2)$ ratio, *i.e.*, in the limit of very poor signal-to-noise data, which in the context of this work correspond to galaxy peculiar velocities, the estimated field approaches zero value. By contrast, the UMV estimator has been specifically designed to not alter the values of the reconstructed field at the locations of the data points, thus avoiding spurious suppression effects. An unbiased estimate of the reconstructed field at any point in space is then obtained by interpolating between the data points, according to the correlation function assumed *a priori*. Like the Wiener filter the new estimator can be used for dynamical reconstructions, *i.e.*, to recover one dynamical field from another measured field, *e.g.*, mass over-density from radial peculiar velocity. These two properties make the UMV estimator a very appealing tool for studying the LSS and CMB.

Here we apply the UMV estimator to reconstruct the density and velocity fields in the local universe from the radial peculiar velocities of the SEcat galaxies. The data points consist of a set of observed radial peculiar velocities, u_i^o, measured at positions $\mathbf{r}_i$ with estimated errors ϵ_i, assumed to be uncorrelated. The observed velocities are thus related to the true 3D underlying velocity field $\mathbf{v}(\mathbf{r})$, or to its radial component u_i , via

$$u_i^o = \mathbf{v}(\mathbf{r}_i) \cdot \mathbf{r}_i + \epsilon_i \equiv u_i + \epsilon_i, \qquad (2)$$

We assume that the peculiar velocity field $\mathbf{v}(\mathbf{r})$ and the density fluctuation field $\delta(\mathbf{r})$ are related via linear gravitational-instability theory, $\delta = f(\Omega_m)^{-1}\nabla\cdot\mathbf{v}$, where $f(\Omega_m) \approx \Omega_m^{0.6}$ and Ω_m is the matter mean density parameter. Under the assumption of a specific theoretical prior for the power spectrum $P(k)$ of the underlying density field, we can write the UMV estimator of the 3D velocity field as,

$$\mathbf{v}^{UMV}(\mathbf{r}) = \langle \mathbf{v}(\mathbf{r}) u_i \rangle \langle u_i u_j \rangle^{-1} u_j^o \qquad (3)$$

and the UMV estimator of the density field as,

$$\delta^{UMV}(\mathbf{r}) = \langle \delta(\mathbf{r}) u_i \rangle \langle u_i u_j \rangle^{-1} u_j^o. \qquad (4)$$

Within the framework of linear theory and assuming that the velocities are drawn from a Gaussian random field, the two-point velocity-velocity and density-velocity correlation matrices (bracketed quantities in eqs. 3 & 4) are readily calculated. Note that the normalization of the power spectrum drops out of the field estimation. The calculation of these matrices is discussed elsewhere (Górski 1988; Zaroubi *et al.*, 1995,1999).

The assumption that linear theory is valid on all scales enables us to choose the resolution as well. In particular it allows us to use two different smoothing kernels for the data and for the recovered fields. In our case no smoothing was applied to the radial velocity data while we choose to reconstruct the density and velocity fields with a finite Gaussian smoothing of radius, R. This choice alters the velocity-velocity and density-velocity correlation functions that appear in the first bracketed terms of the right hand side of eqs. 3 and 4, respectively, by introducing a multiplicative term $\exp[-k^2R^2/2]$ in the model power spectrum.

The amplitude of the reconstructed matter density field given in equation 4 is proportional to $f(\Omega_m)^{-1}$, while that of the density field obtained the PSCz galaxy distribution is proportional to the biasing parameter, b. Therefore, the comparison between these density fields will constrain the value of $\beta = f(\Omega_m)/b$. The velocity field reconstruction from the SEcat data set, however, is independent of Ω_m and b. Hence in the velocity-velocity comparison β enters as a solution of eq. 1 for the PSCz velocities, with the matter density given as the ratio $\delta(PSCz)/b$, where $\delta(PSCz)$ is the PSCz density field.

The error covariance matrix, or variance of the residuals, of the reconstructed density and velocity fields could be calculated theoretically (Zaroubi 2002b). However, in order to give a more complete account of the various errors that enter the calculation (*e.g.*, nonlinear effects) here we choose to calculate the error from mock catalogs (see section 4).

It is interesting to compare the UMV algorithm with other methods of reconstruction used for similar purposes. In the context of mass-density reconstruction from radial peculiar velocities, comparison with the POTENT algorithm (Bertschinger & Dekel 1989, Dekel *et al.*, 1990) is of special interest. The main assumption behind the POTENT algorithm is that the flow field, smoothed on large scale, is derived from a velocity potential. The potential flow assumption is a direct result of linear theory but can also be employed in the quasi-linear regime until the onset of shell crossing, when an extension of eq. 1 applies (Nusser *et al.*, 1991). POTENT could be viewed as a direct inversion method which uses the minimum amount of assumptions, but suffers from the problems of direct deconvolution of very noisy data.

In conclusion, the UMV reconstruction can be regarded as a compromise between the POTENT algorithm, which assumes no regularization but might be unstable to the inversion problem of deconvolving highly noisy data, and the WF algorithm, which takes into account the correlation between the data points and therefore stabilizes the inversion, but constitutes a biased estimator of the underlying field.

3 THE DATASETS

The main dataset used here is the SEcat catalog of galaxy peculiar velocities which results from the merging of 1300 spiral galaxies taken from the SFI catalog (Giovanelli *et al.*, 1997a and 1997b and Haynes *et al.*, 1999a and 1999b) and about 2000 early type galaxies from the ENEAR catalog grouped into ≈ 750 independent objects (da Costa *et al.*, 2000). For each object the radial velocity and inferred distance, corrected for Malmquist bias (Freudling *et al.*, 1995, da Costa *et al.*, 2000), are provided along with the velocity errors that typically amount to $\sim 19\%$ of the galaxy distance.

Merging different peculiar velocity catalogs may result in spurious flows and lead to systematic errors in the reconstruction procedure. However, several pieces of evidence indicate that this is not a serious problem for the SEcat catalog. First of all, both the SFI and ENEAR catalogs are intrinsically homogeneous as they consist of uniform data covering most of the sky. Secondly, as shown in Bernardi *et al.*(2002b), the distance relations independently calibrated in the two sets are consistent with each other and the difference in the Hubble constant deduced from each catalog is $\approx 5\pm10$ km s^{-1}/Mpc. Finally, different statistical analyses carried out using either samples lead to consistent results (*e.g.*, Borgani *et al.*, 2001, da Costa *et al.*, 2001, Nusser *et al.*, 2001, Zaroubi *et al.*, 2001 and references therein). Some of the main characteristics of these samples are summarized in Figure 1. The sky distribution (upper panels), redshifts (central panels) and peculiar velocities (lower panels) of the SFI and ENEAR galaxies are quite similar, especially when accounting for the expected tighter spatial correlation and higher peculiar velocities of the ENEAR early type galaxies. Further evidence for the consistency of the two catalogs will be provided in section 5 (*e.g.*, see Figures 5 and 7) .

The substantial number of galaxies and the large sky coverage (the unobserved region is given by a zone of avoidance of about 15° about the Galactic-plane) allow a dense and uniform sampling of the velocity field. Moreover, since SEcat contains both early and late type galaxies, we can sample both high and low density regions and therefore minimize possible biases that might have affected other analyses based on a single population of objects.

To estimate the β parameter one needs to compare the densities and/or velocities reconstructed from a radial velocity survey with those recovered from a direct probe of the density field, namely a redshift survey. Here we use the models obtained by Branchini *et al.*(1999) from the distribution of IRAS PSCz galaxies under the assumptions of linear biasing and linear theory. The PSCz redshift survey (Saunders *et al.*, 2000) provides the angular positions and redshifts of ~ 15000 IRAS galaxies distributed over almost all sky (the zone of avoidance is about 8°) with a median redshift of 8500 km s^{-1}and is therefore suitable for modeling the density and velocity fields within the same region in which the UMV reconstruction is performed. The typical error associated with the PSCz density and velocity models are significantly smaller than those reconstructed from the SEcat velocities and therefore will be ignored in the subsequent analysis.

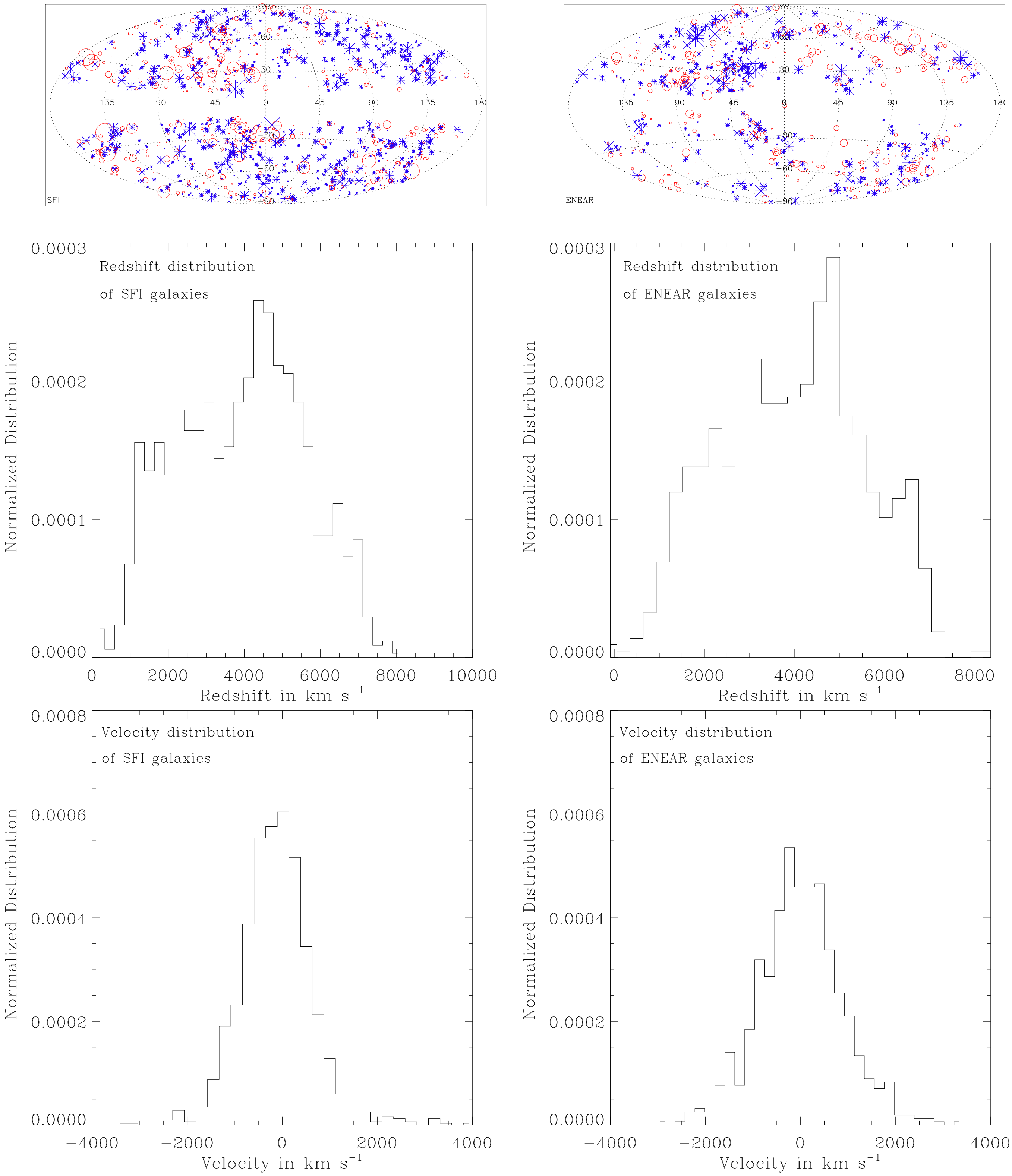

Figure 1. Upper panels: The projected distribution of the SFI (left panel) and ENEAR (right panel) galaxies in galactic coordinates. Crosses indicate positive peculiar velocities, open circles negative; and the size of the symbols is proportional to the amplitude of their peculiar velocity. Middle panels: The redshift distribution of the SFI (left panel) and the ENEAR (right panel) galaxies. Lower panels: The peculiar velocity distribution of the the SFI (left panel) and the ENEAR (right panel) galaxies.

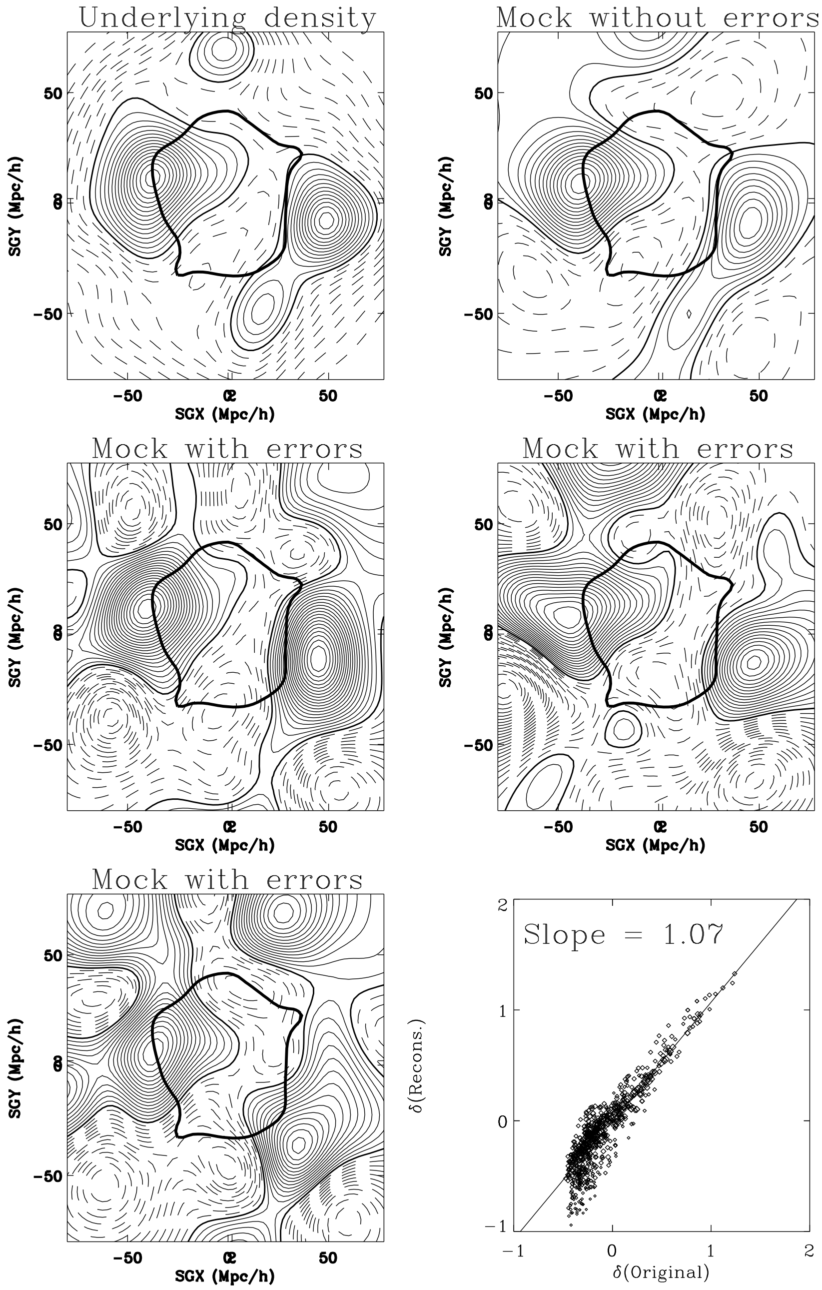

Figure 2. Comparison of the original and reconstructed G12 smoothed density maps on the mock Supergalactic plane. In all panels the solid and dashed line contours denote positive and negative densities respectively. The bold-solid line denotes the zero level density. Contour spacing is 0.1. The very thick solid contour marks the area within which the reconstruction errors are less than 0.2. The upper left hand panel shows the underlying density field of the simulation. The degradation of the original density map towards the edges is spurious and due to the finite size of the smoothing radius. The upper right hand panel shows the reconstructed density from the SEcat mock catalog before adding errors to the distances and velocities. The other three maps show reconstruction from SEcat mock catalogs with realistic noise. The lower right-hand panel shows a typical scatter plot of the density within the comparison region of the original vs. the reconstructed density from one of the mock catalogs (with noise); the points used in the scatter plot are 1/10 randomly sampled from the grid points with estimated reconstruction errors less than 0.2.

4 MOCK CATALOGS AND ERROR ESTIMATES

To test the performance of the method when applied to the SEcat catalog, we construct mock catalogs based on the "Constrained Realization GIF" simulation carried out by Mathis *et al.* (2001). This simulation starts from initial conditions with a smoothed linear density field which matches that derived from the IRAS 1.2 Jy galaxy survey and tracks the formation and evolution

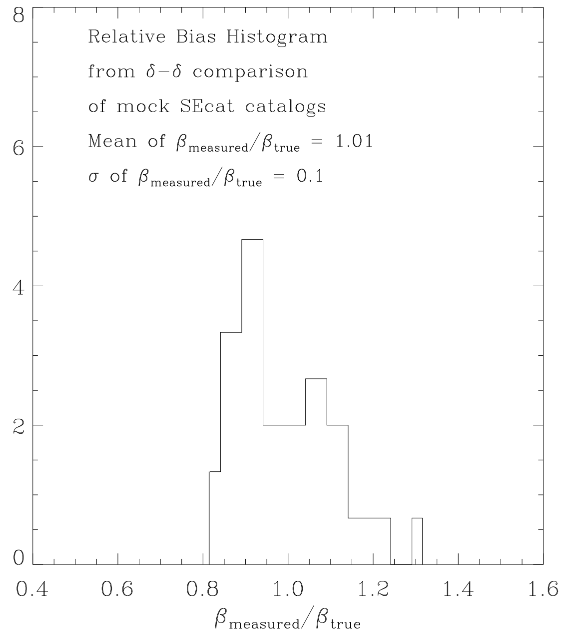

Figure 3. The distribution of the value of β relative to the real one as estimated from 100 mock SEcat catalogs in the density-density comparison (left panel) and velocity-velocity comparison (right panel). The value of the measured β for the catalog without noise is $\beta = 0.99$ and 1.01 for the density and velocity reconstruction respectively.

of all dark matter halos more massive than 10^{11} solar masses out to a distance of 8000 km s^{-1}from the Milky Way up to the current epoch. Galaxies in the original mock catalog are sampled from the N-body simulation at the same position of the real galaxies. Realistic mock catalogs are obtained by assigning errors to the position and velocity of mock galaxies consistent with observations. We construct 100 mock catalogs that differ only in the realization of errors added to the position and velocity of the galaxies. Measured peculiar velocities in the mock catalogs are then obtained after performing a Malmquist bias correction according to the recipe given by Willick *et al.* (1997a).

The density and 3D velocity fields within a region of 60 h^{-1} Mpc , smoothed with a Gaussian filter of radius of 12 h^{-1} Mpc (G12 hereafter), were reconstructed from each of the 100 mock catalogs. Both N-body and reconstructed density and velocity fields were specified on a regular grid with a mesh size of 2.5 h^{-1} Mpc and reconstructed radial velocities at the actual location of the data points were interpolated from the grid. Errors in the reconstructed densities and peculiar radial velocities were estimated from the 100 Monte Carlo realizations.

Note that we do not use the peculiar velocities of the mock galaxy catalogs of Mathis *et al.* (2001). Instead, the peculiar velocities of our mock galaxies are obtained from those of of the dark matter particles. This implies neglecting the so called velocity bias which is expected to be small on the large smoothing scale involved in our reconstruction (see *e.g.*, Carlberg 1994).

Figure 2 shows the quality of the reconstruction from mock SEcat catalogs. The comparison of the original underlying density with the one reconstructed from ideal mock catalogs (i.e. with no velocity errors) in a region in which the estimated reconstruction errors are ≤ 0.2 shows that their relation is well described by a linear function with a slope of 0.99 (see below). This demonstrates that the sampling density of the data is sufficiently high. Note that the density of the original catalogs close to the boundaries is suppressed. This effect is due to finite smoothing length and does not affect our analyses which are carried out to distances well within the edges of the mock catalogs. The other three maps shown in Figure 2 were randomly chosen from the 100 mock catalogs reconstructions. The similarity of the reconstructions are quite evident, especially within the region of small reconstruction errors. The scatter plot quantifies this similarity for one of the mock catalogs. An inspection of the radial peculiar velocity reconstruction shows a similar quality of results.

Left panel of Figure 3 shows β as estimated by applying the UMV reconstruction algorithm to the mock catalogs and comparing the results with the true density and velocity fields using the following χ^2 statistic,

$$\chi^2 = \frac{1}{N} \sum_{\sigma_\delta \leq 0.2} \frac{[\delta_i(\mathrm{Mock}) - \beta(\delta_i(\mathrm{Original}) + \Delta\delta)]^2}{\sigma_\delta^2}, \quad (5)$$

where δ_i(Mock) is the density as reconstructed from the mock SEcat catalog; δ_i(Original) is the original density, and σ_δ is the rms difference between the reconstructed densities from the mock catalogs and the original density at the same point in space. The free parameters here are β and the offset between the two fields, $\Delta\delta$. The latter is introduced to account for the uncertainty in defining the mean density of the sample. This offset is expected to be zero in the mock catalog analyses and can only be non trivial when comparing the density fields obtained from two different catalogs (such as the true SEcat and PSCz catalogs). The density-density comparison is carried out over N points, randomly selected from the $10 \times N$ at which the estimated errors are less than 0.2. The expected values of β and $\Delta\delta$ are unity * and zero respectively. The left panel of Figure 3 shows that the estimated β is unbiased and has a mean and variance of 1.01 and 0.1, respectively. The estimated value of $\Delta\delta$ is -0.03 and has a variance of 0.04.

Naturally, most of the grid points used in the comparison – even after diluting their number by a factor of 10 – do not have statistically independent errors. Therefore, the statistic used in eq. 5 is not optimal. In other words, had the data points used in eq. 5 been independent then it would indeed be sufficient to use the likelihood contours obtained after minimizing the χ^2 statistic to determine the uncertainty of the results. On the other hand, had these data point been totally dependent with about only one degree-of-freedom then the error obtained from the likelihood analysis would be a gross underestimate unless the *cosmic-variance like* errors over the whole sample were also estimated and taken into account. Normally, the best way to go about estimating the scatter in the results that takes into account the partial dependency of the data points and their limited number is to perform a very large number of Monte Carlo simulations – of the order of 10^2-10^3 – to ensure an accuracy of a few percent. This however is very time consuming and not feasible with available computer resources. Therefore, we choose to assume that the errors are totally independent and use eq. 5 in order to estimate the most likely β and $\Delta\delta$ values. The uncertainties are then determined by adding in quadrature the likelihood errors and the uncertainties obtained from the scatter in the value of β, obtained from the 100 mock catalogs. In our sample, these two sources of errors are not totally independent. However, by treating them as such allows one to obtain a conservative upper limit on the errors.

In order to validate our approach it is important to estimate the number of degrees-of-freedom in the sample, $\mathcal{N}_{dof}$. In the case of no noise correlations except those introduced by the G12 smoothing, $\mathcal{N}_{dof} \approx 23$ in a sphere of radius $50h^{-1}$ Mpc , which represents the ratio of the total volume to the effective volume of the G12 filter. In the case at hand, however, the calculation is more subtle and involves numerical estimation of the noise correlation matrix from the 100 Monte-Carlo simulations at each grid point with error less than 0.2 and finding its eigenvalues. Then $\mathcal{N}_{dof}$ is identified with the number of significant eigenvalues of this matrix (see Zaroubi *et al.*, 1995 for the treatment of a similar problem). Specifically, $\mathcal{N}_{dof}$ is the number of the highest eigenvalues that account for 95% of the variance, found here to be about 20 (see Figure 4). The eigenvalues of most of the remaining eigenmodes drop by orders of magnitude. This number, 20, reflects the number of effective smoothing volumes and the additional correlation introduced by the UMV filter.

* The density reconstruction is performed hereafter normalizing with the correct value of $f(\Omega_m = 0.3)$ therefore the comparison is expected to yield a bias parameter of unity.

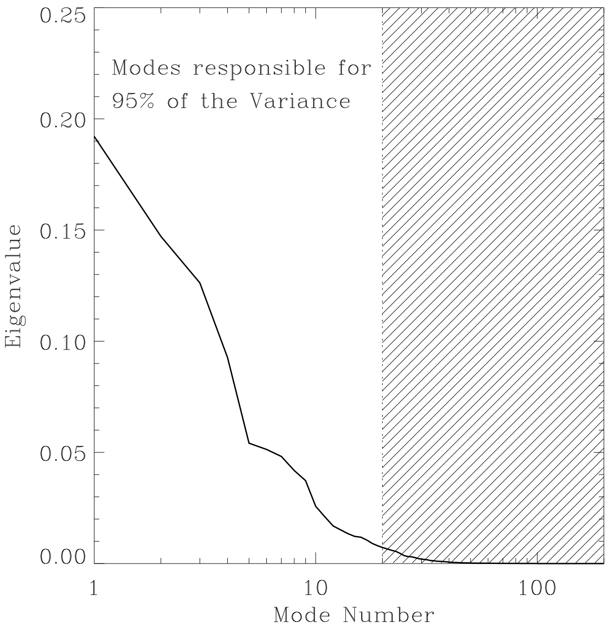

Figure 4. The sorted spectrum of the eigenvalues of the noise correlation matrix. The modes that account for 95% of the variance are the first 20.

The same strategy is adopted for the velocity-velocity comparison. In this case the statistic of choice is,

$$\chi^2 = \frac{1}{N_{data}} \sum_{Data Points} \frac{(u_i(\mathrm{Mock}) - \beta u_i(\mathrm{Original}) - \Delta H_\circ r_i)^2}{\sigma_v^2}, \quad (6)$$

where u(Mock) and u(Original) are the radial velocities of the mock and original data respectively. N_{data} is the number of data points, $\Delta H_\circ$ is the offset in Hubble constant, r_i is the distance of the data point i and σ_v is the uncertainty in the radial velocity as obtained from the mock catalogs. The velocity-velocity comparison is carried out over the radial velocities at the location of the data points. The argument regarding the χ^2 statistic used here is identical to the one discussed for the density-density comparison. Right panel of Figure 3 shows that the mean and variance of β in this case are 0.97 and 0.12, respectively. The offset in Hubble constant is $0.2 \pm 0.5 \mathrm{km\,s^{-1}/Mpc}$. The value of β estimated from the noise-free mock catalog is 1.03. From the noise correlation matrix we estimate $\mathcal{N}_{dof} \approx 17$. The lower value obtained here reflects the longer range of velocity correlation.

Willick and Strauss (1997b) have suggested in their VELMOD method to estimate β from velocity data by minimizing residuals and their correlations simultanously with minimizing the likelihood, therefore, allowing the calibration of the distance estimator while evaluating β. Whereas in our analysis we assume that the distance indicator has been calibrated independently.

The analysis of the mock catalogs shows that the

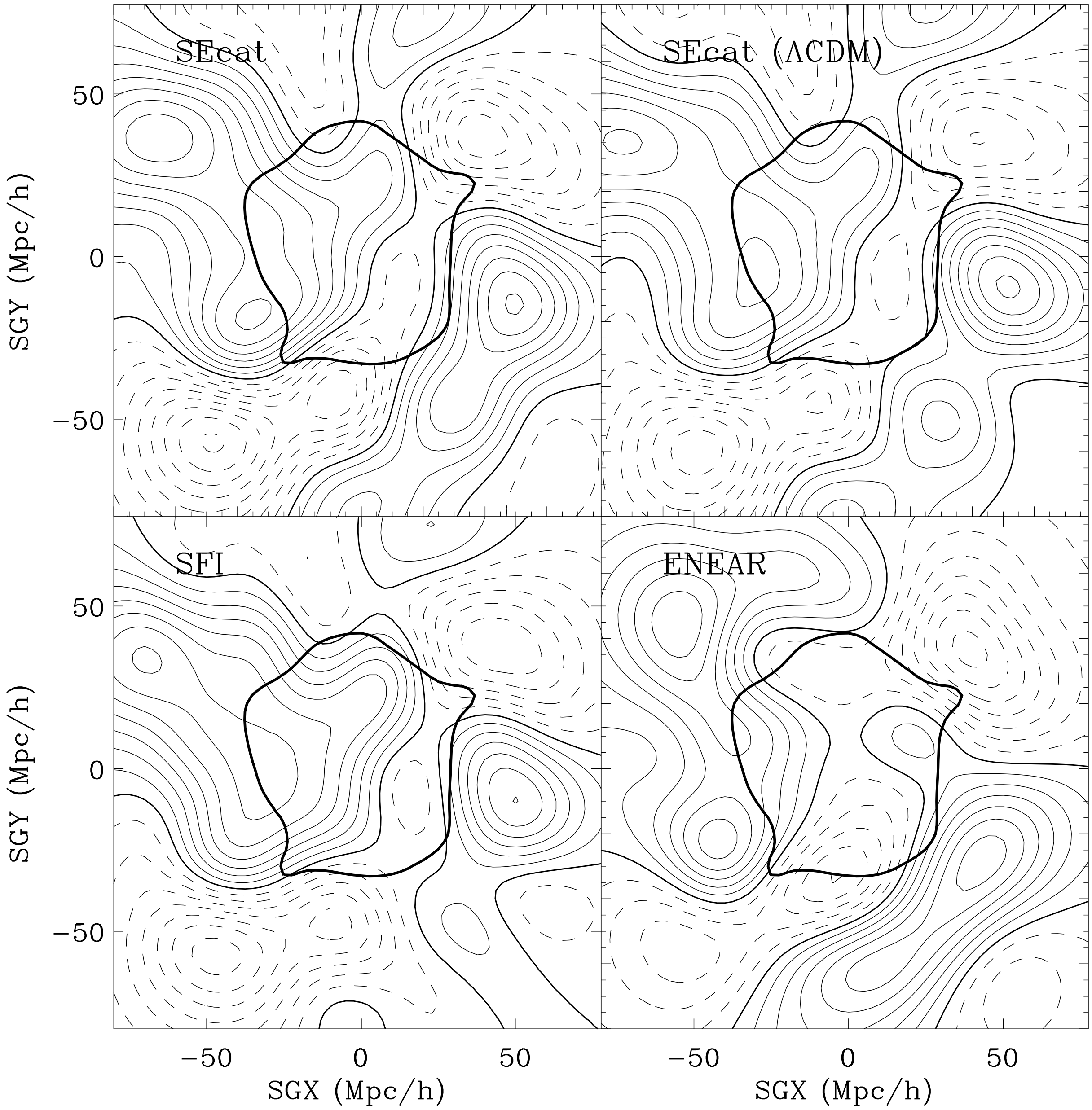

Figure 5. The G12-smoothed, overdensity field on the Supergalactic plane reconstructed from SEcat, ENEAR and SFI catalogs. All reconstructions assume a Standard CDM power spectrum apart from the map shown in the upper right panel for which we have assumed an $\Omega_m = 0.3$ ΛCDM power spectrum. The very thick solid contour marks the area within which the reconstruction errors from the SEcat catalog are less than 0.2. In all panels the solid and dashed line contours denote positive and negative densities respectively. The bold-solid line denote the zero level density. Contour spacing is 0.1.

UMV returns an unbiased estimate of the underlying density and velocity fields from which β can be determined with an accuracy of 10%. This error is purely random and takes into account the scatter of the best fit β value but does not include the scatter in the estimation of each individual value of β which will be added later.

This analysis could in principle be extended to smaller smoothing kernels. Besides the validity of linear theory required in this analysis, the density of the sky coverage might be also an important factor since an insufficient coverage might result in the dominance if noise in most of the sampling volume. An extension of the analysis to a $900\,\mathrm{km\,s^{-1}}$ smoothing kernel (hereafter, G9) shows that the results of the analysis are consistent with the G12 comparison but the uncertainties are much higher especially in the density-density comparison. Therefore, the smoothing kernel used in the rest of the paper is G12.

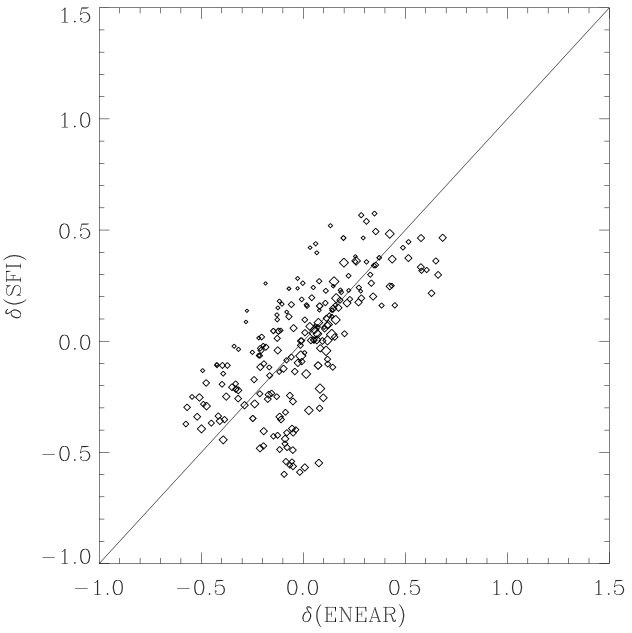

Figure 6. A quantitative comparison between the SFI and ENEAR G12-smoothed. reconstructed densities. The densities of both SFI and ENEAR were reconstructed on a grid with mesh size of 2.5 h^{-1} Mpc The densities shown in the scatter plot are from gridpoints randomly selected with a rate 1/10 from those with reconstruction errors less than 0.2. The size of the symbols is inversely proportional to their errors. The solid line with a slope of unity has been drawn to guide the eye. The agreement between the two reconstructions is very good except for a small number of points with $\delta(\mathrm{SFI}) \approx -0.5$ and $\delta(\mathrm{ENEAR}) \approx 0$.

5 THE FIELD-FIELD COMPARISON

Here we apply the UMV estimator to the true SEcat catalog to obtain the G12-smoothed density and velocity fields assuming a power spectrum of a flat CDM model with $h = 0.5$, $\Omega_m = 1.0$ as a prior. These cosmological parameters determine the shape of the power spectrum only and are not used anywhere else in the analysis. Unlike the outcome of the Wiener Filter estimator, the UMV-reconstructed density and velocity fields, for dense sampling of the sky, are unbiased and therefore can be used for quantitative comparisons.

Before proceeding further in analysing the SEcat catalog, it is useful to compare the reconstruction we obtain from the SEcat peculiar velocities with those obtained from the SFI and ENEAR separately in order to support our claim regarding the consistency the latter two. Figures 5, 6, and 7 show the similarity between the density and velocity reconstruction of the three catalogs. We also note that, as demonstrated in the same figures, changing the power spectrum prior has a very small effect on the results.

The reconstruction of the PSCz model density and velocity fields, performed according to Branchini *et al.*(1999), requires an input value for β; here we use $\beta = 0.5$. However, the dependence of the reconstructed density field on the input β is very weak while the model peculiar velocity roughly scales linearly with β. This dependence results in a systematic error of about 2% on the final result. Model PSCz velocities were obtained at the reconstructed positions of PSCz galaxies in real space. Both masses and velocities were then smoothed with a G12 filter to obtain the PSCz density and velocity fields on a regular grid with a mesh size of 2.5 h^{-1} Mpc Finally, G12-smoothed PSCz velocities have been interpolated at the positions of SEcat galaxies. Here we compare the SEcat fields with those of the PSCz, all smoothed with a G12 filter. The errors on the model density and velocity fields are much smaller than the SEcat ones (see Branchini *et al.*, 1999) and will be ignored in the following comparisons.

5.1 The density-density comparison

The left panels of Figure 8 show the G12 smoothed UMV reconstructed density field map in three planes at different Supergalactic Z (the central plane refers to Z=0, i.e. the Supergalactic plane) obtained from the SEcat peculiar velocities within a box of 160 h^{-1} Mpc aside, centered around the Local Group position. The main features of our local universe are easily identified in the UMV map on the Supergalactic plane , including the Great Attractor on the left and the Perseus-Pisces supercluster in the lower right. There is also a hint of the Coma cluster, which lies just outside the sample, in the upper part on the map. Similar features also characterize the PSCz density map shown on the right-hand panel of Figure 8. This map, obtained by Branchini *et al.*(1999), has the same smoothing (G12) and shows the same region of the universe.

A quantitative density-density comparison is carried out using the following χ^2 statistic:

$$\chi^2 = \frac{1}{N} \sum_{\sigma_\delta \leq 0.2} \frac{[\delta_i(\mathrm{SEcat}) - \beta(\delta_i(\mathrm{PSC}z) + \Delta\delta)]^2}{\sigma_\delta^2}, \qquad (7)$$

where $\delta_i(SEcat)$ and $\delta_i(\mathrm{PSC}z)$ are the SEcat and PSCz densities respectively, $\Delta\delta$ is the offset in the mean density and σ_δ are the density reconstruction error at each point as estimated from Monte Carlo realizations of mock-SEcat catalogs. The best fit β and $\Delta\delta$ parameters are those that minimize Eq. 7. Here again the sum is over gridpoints randomly sampled at a rate of 1/10 from those that have errors less than 0.2. As in the mock catalogs analysis, here $\mathcal{N}_{dof} \approx 20$.

In previous density-density comparisons (*e.g.*, Sigad *et al.*, 1998) the authors chose to minimize the χ^2 with respect to $1/\beta$ instead of β directly. Since the main source of errors in our analysis are the uncertainties in the measured galaxy velocities, adopting a direct or inverse β minimization of the χ^2 statistic does not make much difference. Here we choose to minimize with respect to β.

Each point in the scatter plot displayed in the lower panel of Figure 9 shows the comparison between the SEcat and PSCz overdensities, measured at the same locations. The comparison is restricted to the 1/10 randomly chosen points with $\sigma_\delta < 0.2$. The slope of the solid line gives $\beta = 0.57$.

A zero-point offset $\Delta\delta = 0.18$ is also detected. We

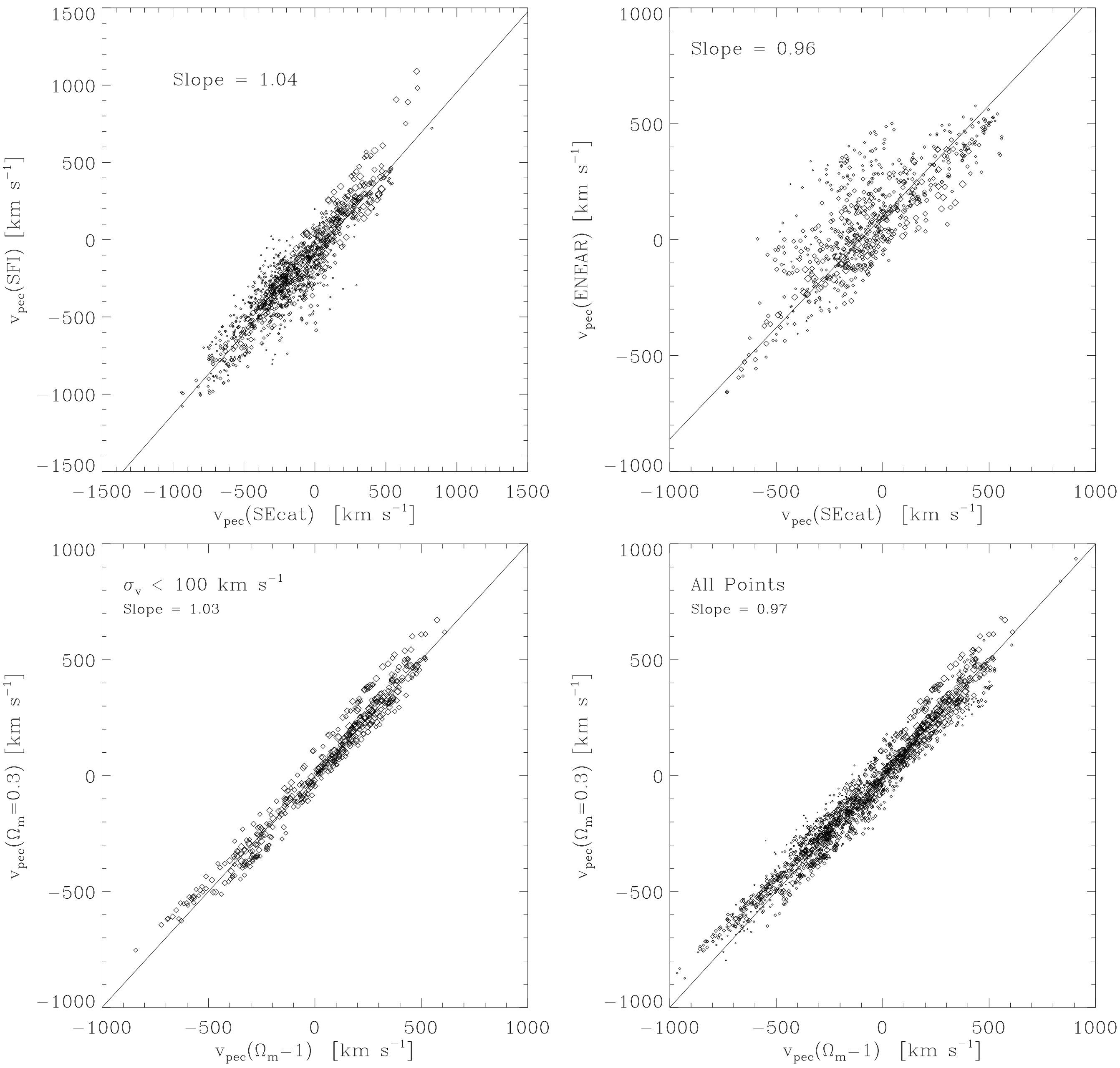

Figure 7. The top left panel compares the G12-smoothed peculiar radial velocities reconstructed from the SFI catalog *versus* those reconstructed from the SEcat catalog, both measured at the location of the SFI data. The solid line represents the best linear fit to the scatter plot. The top-right panel compares ENEAR and SEcat reconstructed velocities at the positions of the ENEAR galaxies. The two bottom panels refer to the SEcat catalog only and compare velocities reconstructed assuming a ΛCDM model (Y axis) with those reconstructed from the standard CDM model (X axis). Peculiar velocities shown in the bottom-left panel refer to points with estimated error less than $100/kms$. All points are considered in the bottom-right panel. In all plots the size of the symbols is inversely proportional to their errors.

interpret it as a mismatch in the average density in the two samples. The mismatch in the mean fields is caused by the PSCz density field which was found to be systematically larger than the IRAS 1.2Jy density field within a $60\,h^{-1}$ Mpc sphere (Teodoro *et al.*, 2000). This mismatch arises due to the incompleteness of the PSCz catalog at low fluxes.

The upper panel in Figure 9 shows the 1, 2 and 3-σ likelihood contours in the β–$\Delta\delta$ plane obtained from Eq. 7. The marginalization of this distribution with respect to $\Delta\delta$ gives the error estimate on the value of β which is of the order of 0.08 ($\approx$ 15%). Adding to this error the error estimated from the distribution of bias shown in Figure 3 one obtains $\beta = 0.57^{+0.11}_{-0.13}$.

To estimate the goodness-of-fit of the parameters obtained from the χ^2 analysis, we calculate the distribution of the residuals,

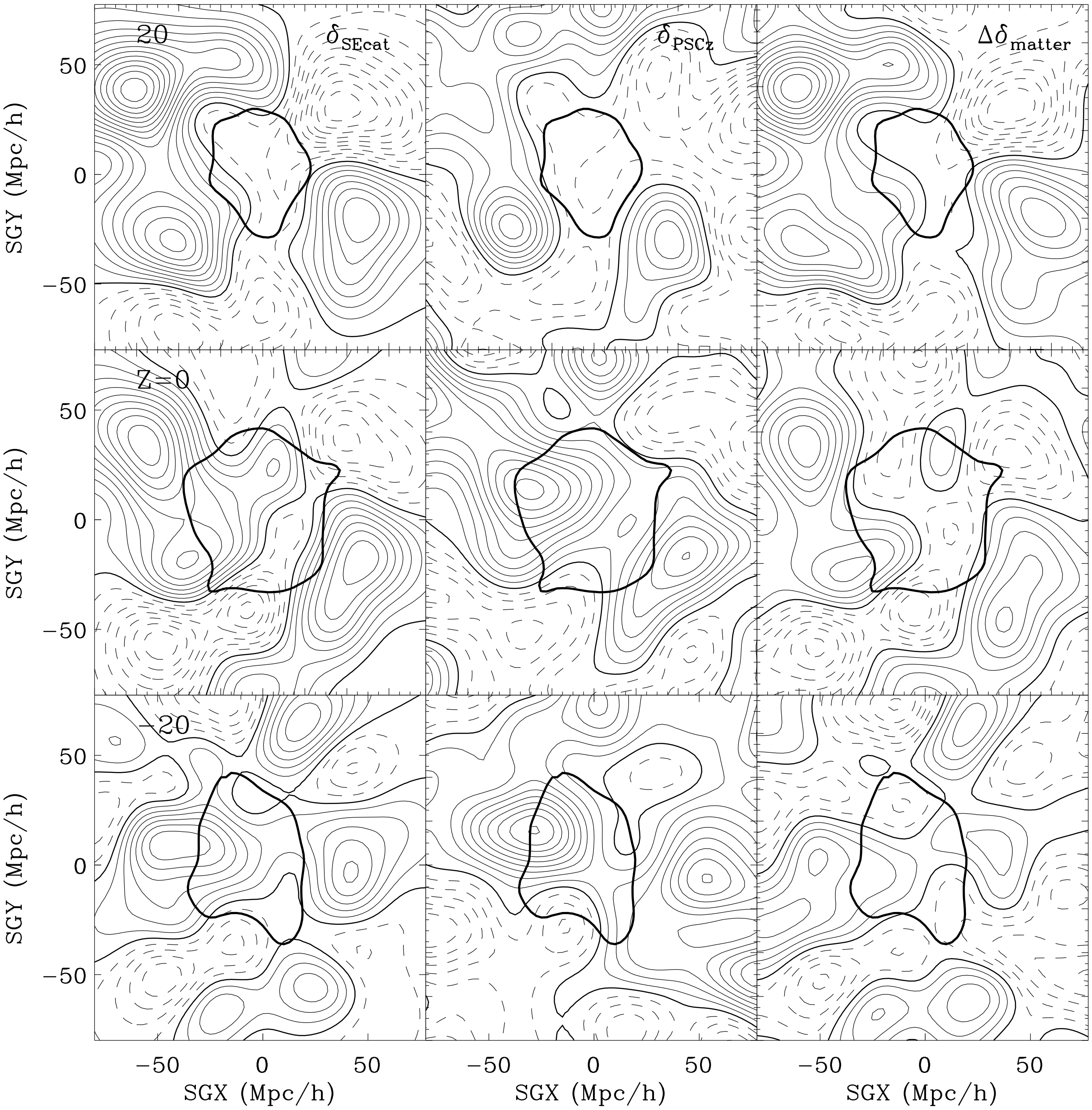

Figure 8. The panels on the left show the maps of the G12-smoothed, density fluctuations UMV-reconstructed from the SEcat catalog. The central panels show the same field reconstructed from the distribution of PSCz galaxies. The maps of density residuals, computed for $\beta = 0.57$ $\Delta\delta = \delta_{SEcat} - \delta_{PSCz}$ are shown on the right-hand panels. The central maps show the density fields on the Supergalactic planes. The maps on the upper and lower panels show the density fields at Supergalactic Z= $\pm 20^{-1}$ Mpc, respectively. The very thick contour marks the boundaries of the volume within which the estimated reconstructed error is ≤ 0.2; the volume used for comparisons.

$$\xi = \frac{\delta(\mathrm{SEcat}) - \beta_{ML}\big(\delta(\mathrm{PSC}z) - \Delta\delta\big)}{\sigma_\delta}, \tag{8}$$

where β_{ML} is the best fit β parameter and σ_δ is the error on the UMV estimated density field. If the model correctly describes the data, this distribution should be Gaussian with a rms of unity. The histogram of ξ is shown in Figure 10 along with the best fitting Gaussian distribution (dashed line), whose rms is almost unity.

To check the robustness of the result we have repeated the density-density comparison by drawing three additional samples defined at different error thresholds and re-estimating the values of β. Adopting the thresh-

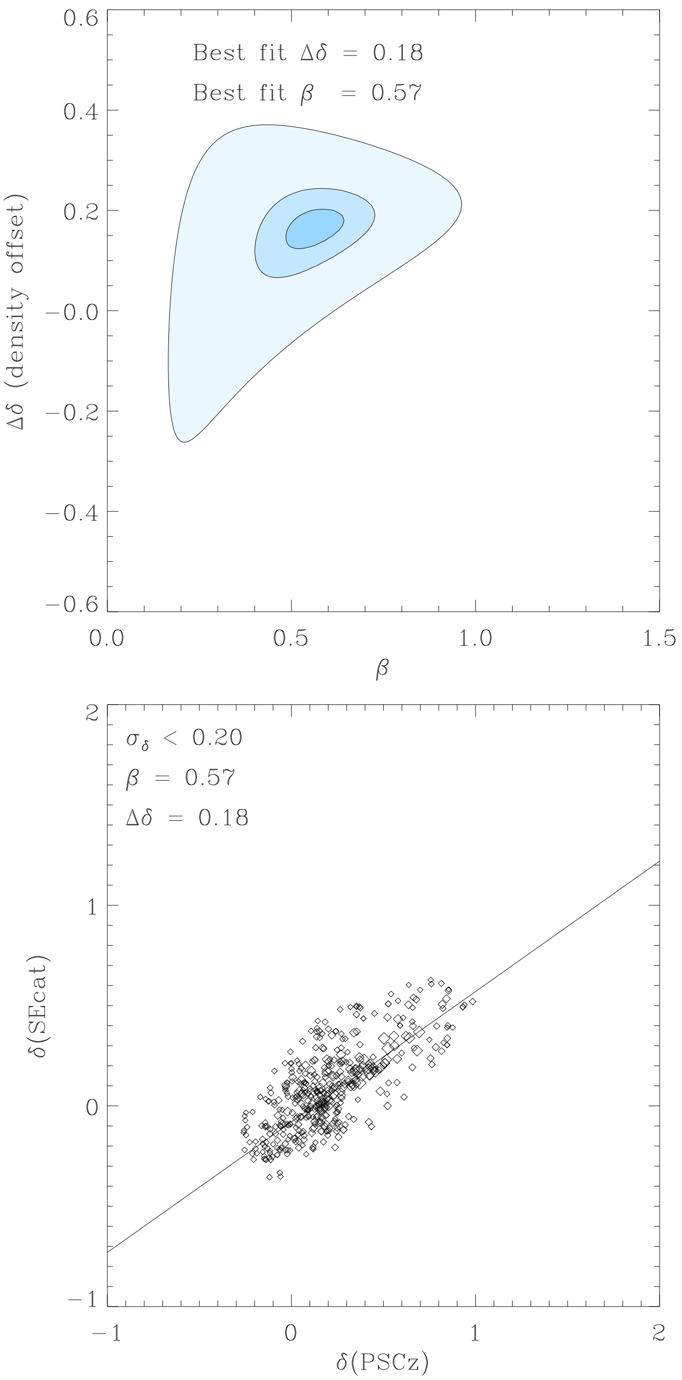

Figure 9. The upper panel shows the 1, 2 and 3 σ likelihood contours from the $\delta-\delta$ comparison in the $\beta-\Delta\delta$ plane. The scatter plot in the bottom panel compares the G12-smoothed UMV-reconstructed SEcat density field to the PSCz density fluctuations, both measured at grid points 1/10 randomly selected from those with reconstruction errors $\sigma_\delta < 0.2$ The size of the symbols is inversely proportional to their errors.

olds $\sigma_\delta = 0.3, 0.4$ and 0.5 we obtained $\beta = 0.54, 0.53$ and 0.52, respectively.

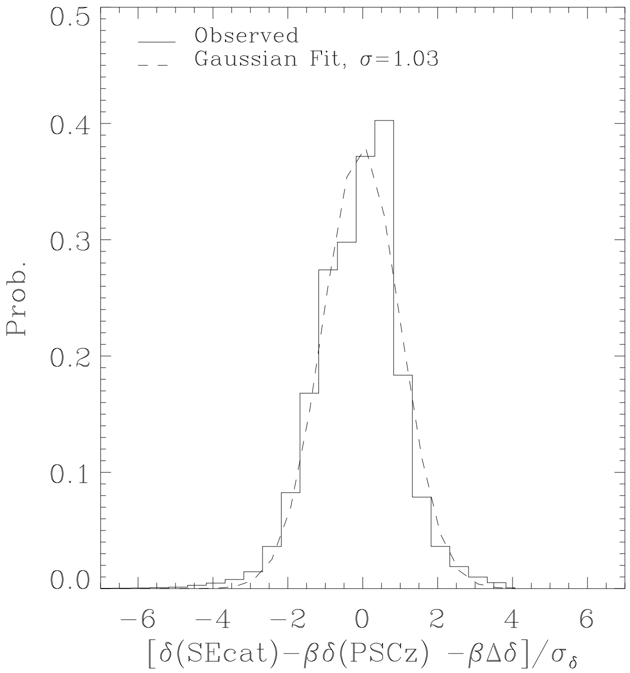

Figure 10. The histogram represents the distribution of the normalized residuals (ξ). The dashed line shows its best Gaussian fit with a rms value of $\sigma = 1.03$, as indicated in the plot.

5.2 The velocity-velocity comparison

The UMV-velocity reconstruction procedure is different from the one used for the density. SEcat radial velocities are calculated at the location of the data points from the reconstructed 3D velcoty which has been homogeneously smoothed with a G12 filter through the UMV operation. Here we compare these velocities to the G12-smoothed model PSCz velocity sampled at the same locations. Therefore, unlike for the density-density case, the number of points used in the velocity-velocity comparison is determined directly by the number of points in the SEcat catalog. This comparison is carried out using the following χ^2 statistic

$$\chi^2 = \frac{1}{N} \sum_{Data Points} \frac{(u_i(\mathrm{SEcat}) - \beta u_i(\mathrm{PSC}z) - \Delta H_\circ r_i)^2}{\sigma_v^2}, \quad (9)$$

where u denotes the G12-smoothed radial velocities, $\Delta H_\circ$ is a local perturbation to the Hubble constant, r_i is the radial distance of the point i and σ_v is the error in the UMV reconstruction. The resulting velocity-velocity scatter-plot is shown in the lower panel of Figure 11 along with the best fitting line. Here the slope of the line constitutes an estimate of $\beta = 0.51$. The zero-point mismatch, $\Delta H_\circ$ represents a spurious "breathing-mode" which is to be expected given the average density mismatch found in the density-density comparison. A perturbation of $\Delta H_\circ = 1.5 \mathrm{km\,s^{-1}}/Mpc$ was found and its associated spurious radial motion, $\Delta H_\circ r$, was subtracted from the PSCz peculiar velocities shown in the lower panel of Figure 11. The two zero-points are found here to be consistent with the prediction of the linear theory relation:

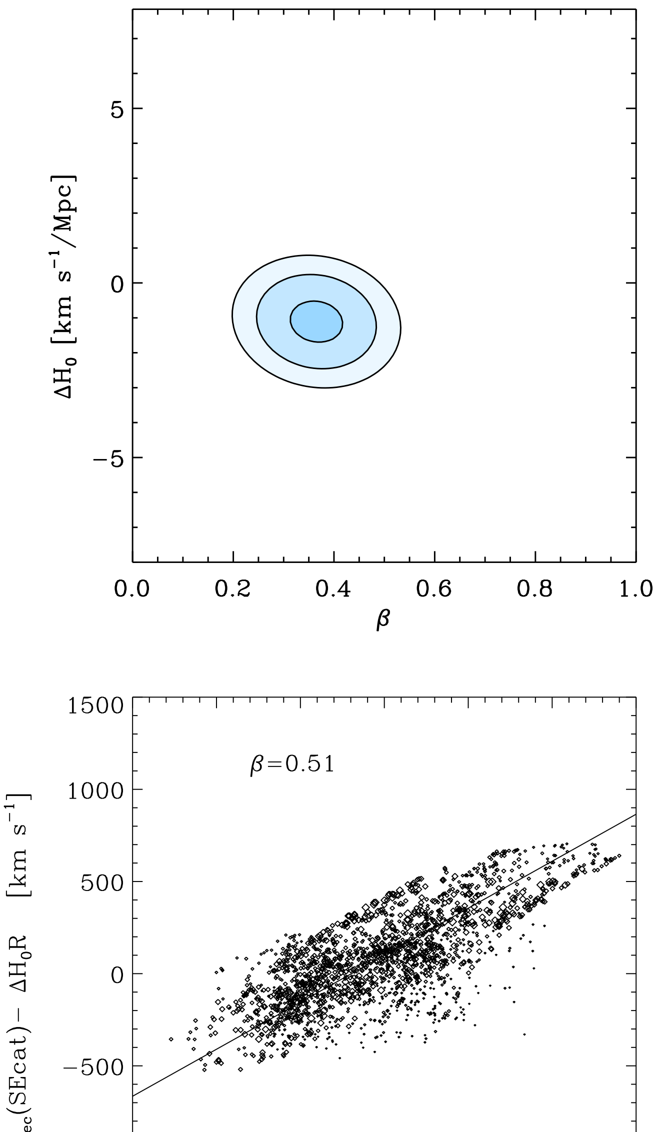

Figure 11. Upper Panel: The 1,2, and 3 σ likelihood contours in the $\beta - \Delta H_\circ$ plane from the SEcat *vs.* PSCz velocity-velocity comparison. In the bottom panel the G12-smoothed, reconstructed SEcat and PSCz are compared at the locations of the SEcat data points. The size of the symbols is inversely proportional to the reconstruction errors

$$\Delta H_\circ = -\frac{\Omega_m^{0.6}}{3} \Delta\delta(< r) H_\circ, \tag{10}$$

where $\Delta\delta(< r)$ is the mean-density mismatch within a radius r.

The upper panel of Figure 11 shows the likelihood contours in the $\beta - \Delta H_\circ$ plane obtained by calculating the χ^2 distribution given in Eq. 9. From the iso-probability contours of 1,2 and 3-σ levels shown in the upper panel of Figure 11 we obtain that $\beta = 0.51 \pm 0.06$, fully consistent with the estimate of β from the density-density comparison. The uncertainty here is a combination of the errors estimated from Figures 3 and 11.

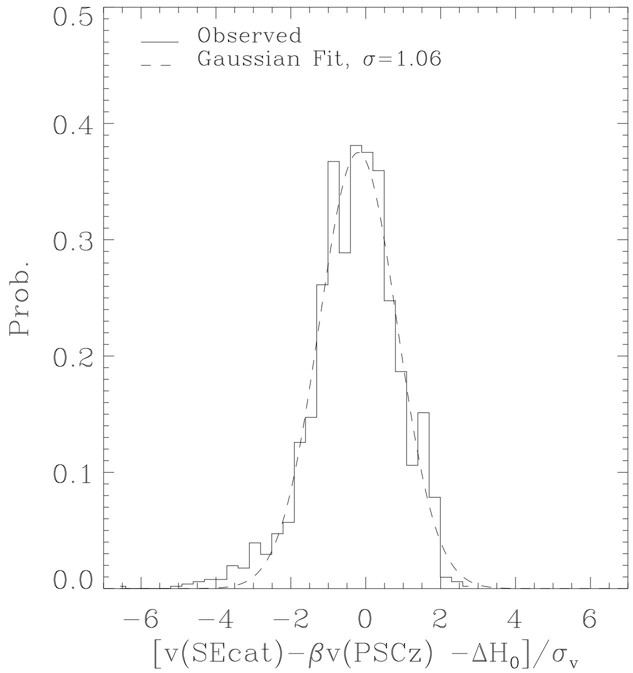

Figure 12. The distribution of the residuals from the velocity velocity comparison. The dashed line shows a Gaussian fit with $\sigma = 1.06$.

The analysis of the velocity-velocity residuals is performed similarly to those of the density-density residuals. The resulting distribution of the residuals is shown in Figure 12 as a histogram. Here again the rms value found for the best fit Gaussian distribution (dashed line) is very close to unity indicating the adequacy of the PSCz velocity model with $\beta = 0.51$. The slight excess in the positive tail of the histogram is due to presence of few outliers,easily identified below the best fitting line in the lower panel of Figure 11. The size of those points clearly indicates that they have very large errors, which would exclude them from any reasonable noise data-cut analysis. Moreover, their number is very small and therefore their influence on the final result is negligible.

6 SUMMARY AND DISCUSSION

In this paper we have applied the new UMV estimator to recover the density and velocity fields in the local universe from the SEcat catalog of galaxy peculiar velocities. In order to obtain the so called β parameter, these fields have been compared with those modeled from the spatial distribution of IRAS PSCz galaxies assuming linear theory and biasing. Previous estimates of β from density-density comparisons, mainly based on the POTENT algorithm (Bertschinger & Dekel 1989, Dekel *et al.*, 1990), have yielded a large value ($\beta \approx 1$ *cf.*, Sigad *et al.*, 1998), inconsistent with the smaller values (≈ 0.5) independently obtained from all recent velocity-velocity VELMOD (Willick *et al.*, 1996, Willick & Strauss 1998, and Branchini *et al.*, 2001) and ITF (da Costa *et al.*, 1998 and Nusser *et al.*, 2000) comparisons.

For the first time the UMV method provides a common methodological framework in which to perform velocity-velocity and density-density comparisons. The velocity-velocity comparison yields a value of β consistent with that measured in the VELMOD and ITF analyses. However, the value of the same parameter obtained from our density-density comparison is significantly smaller than those obtained from the POTENT analyses (*cf.*, Sigad *et al.*, 1998). The β parameters from both $v-v$ and $\delta-\delta$ comparisons presented here are in agreement, yielding a $\beta \approx 0.55$ with an estimated error of the order of 0.1.

In contrast with the POTENT algorithm, the new UMV method reconstructs the density field from peculiar velocities while taking into account their underlying correlation properties. The regularization aspect of the UMV estimator significantly improves the stability of the inversion, which is especially important given the low signal-to-noise ratio of peculiar velocity data. The regularization obtained by this method is very similar to the one provided by the Wiener filter method (Zaroubi *et al.*, 1999). However, the UMV has been designed to provide an unbiased estimator of the underlying signal, a property that the Wiener filtering method lacks. These two aspects make the UMV estimator a very useful tool for reconstruction from peculiar velocity data.

In our error analysis we have shown that for the best fit value of β the density and velocity residuals are normally distributed. This indicates that the PSCz density and velocity fields constitute an adequate model for those reconstructed with the UMV estimator. The fields only differ by a monopole term, corresponding to a mismatch in the the mean density within $60\,h^{-1}$ Mpc which is caused by the known incompleteness of the PSCz catalog at faint fluxes (Teodoro *et al.*, 2000). This also implies that the effect of the nonlinear dynamics and amount of nonlinear and stochastic biasing on the scales involved in our analysis is negligible relative to the measured peculiar velocity errors.

The results presented in this paper are quite encouraging since for the first time the two ways of estimating the value of the β parameter give a consistent result. Our results also suggest that the UMV estimator is a promising tool for the problem of reconstructing the dynamical fields from the observed radial peculiar velocities and therefore could be applied to other datasets. In particular, to reconstruct the large scale structure from the incoming large and uniform surveys that will provide both the spatial distribution and peculiar velocities of extra-galactic objects *e.g.*, SDSS and large cluster surveys with kinematic Sunyaev-Zel'dovich measurements.

Our present density-density comparison results are in marked contrast to those obtained by earlier work, including those from the recent analysis of the Mark III catalog using the POTENT method (*e.g.*, Sigad *et al.*, 1998), and raises the question on the origin of this discrepancy. The Mark III catalog, as shown for example by Davis *et al.* (1996) and more recently by Courteau *et al.* (2000), suffers from systematic calibration errors that would cause a systematic error in the estimation of β. However, these errors are not expected to overestimate the value of β by more than a factor of two. An application of the UMV method to the Mark III catalog shows that the obtained values of β are somewhat higher than those obtained from the SEcat catalogs by $0.1-0.2$. Moreover, the $v-v$-like VELMOD analysis yield consistent values of β when applied to Mark III and SFI datasets (Willick *et al.*, 1997b, Willick and Strauss 1998, Branchini *et al.*, 2001). Based on these arguments, one could speculate that the most likely explanation to the inconsistent results is a conspiracy of both the systematics errors in Mark III and some noise-driven inversion instability in the POTENT reconstructions.

ACKNOWLEDGMENTS

The authors would like to thank the referee, Michael Strauss, for his very useful comments and suggestions, and Anthony J. Banday for critical reading of the manuscript. LdC and SZ would also like to thank the entire ENEAR team. YH and SZ thanks the Universitá di Roma Tre. EB and YH thank the Max Planck Institut für Astrophysik for their hospitality while part of this work was done. YH has been supported in part by the Israel Science Foundation (103/98).

REFERENCES

Baker J., Davis M., Strauss M., Lahav O., Santiago B. 1998 ApJ, 508, 6

Berlind A.A., Narayanan V.K., Weinberg David H., 2001, ApJ, 549, 688

Bernardi, M., Alonso, M. V., da Costa, L. N., Willmer, C. N. A., Wegner, G., Pellegrini, P. S., Rité, C., Maia, M. A. G.,2002, AJ, 123, 2990.

Bertschinger E., Dekel A. 1989, ApJ, 336, L5

Blanton, M., Cen, R., Ostriker, J.P., Strauss, M.A., 1999, ApJ,522, 590.

Branchini E., Teodoro L., Frenk C., Schmoldt I., Efstathiou G., White S., Saunders W., Sutherland W., Rowan-Robinson M., Keeble O., Tadros H., Maddox S., Oliver S., 1999, MNRAS, 308, 1

Branchini E., Freudling, W. da Costa L., Frenk C., Giovanelli R., Haynes M., Salzer J., Wegner G., Zehavi I. 2001, MNRAS, *in press*

Borgani, S., Bernardi, M., da Costa, L.N., Wegner, G., Alonso, M. V., Willmer, C. N. A., Pellegrini, P. S., Maia, M. A. G., 2000, ApJ, 537, L1

Carlberg R. G., 1994, ApJ, 434, L51

da Costa L., Nusser, A., Freudling W., Giovanelli R., Haynes M., Salzer J. Wegner G. 1998, MNRAS, 299, 452

da Costa L. Bernardi M., Alonso M., Wegner G., Willmer C., Pellegrini P., Rité C., Maia M. 2000, AJ, 120, 95

da Costa, L. N., Bernardi, M., Alonso, M. V., Wegner, G., Willmer, C. N. A., Pellegrini, P. S., Maia, M. A. G., Zaroubi, S., 2000, ApJ, 537, L81

Courteau, S., Willick, J. A., Strauss, M. A., Schlegel, D., Postman, M., 2000, ApJ, 544, 636.

Davis, M., Nusser, A., Willick, J.A., 1996, ApJ, 473, 22.

Dekel A., Lahav O. 1999, ApJ, 520, 24

Dekel A., Bertschinger E. & Faber S.M. 1990, ApJ, 364, 349

Freudling W., da Costa L., Wegner G., Giovanelli R., Haynes M., Salzer J. 1995, AJ, 110, 1995

Freudling W, Zehavi, I, da Costa L. N., Dekel A., Eldar A., Giovanelli R., Haynes M. P., Salzer, J. J., Wegner G., Zaroubi S., 1999, ApJ 523, 1.
Giovanelli R., Haynes M., Herter T., Vogt N., Wegner G., Salzer J., da Costa L., Freudling W. 1997a, AJ, 113, 22
Giovanelli R., Haynes M., Herter T., Vogt N., da Costa L., Freudling W., Salzer J., Wegner G. 1997b, AJ, 113, 53
Gorski, K.M., 1988, ApJ, 332, L7
Haynes, M., Giovanelli R., Salzer J., Wegner, G., Freudling W., da Costa L., Herter T., Vogt N. AJ, 1999a, 117, 1668
Haynes, M., Giovanelli R., Chamaraux P., da Costa L., Freudling W., Salzer J., Wegner, G. AJ, 1999b, 117, 2039
Mathis, H., Lemson, L., Springel, V., Kauffmann, G., White, S.D.M., Eldar, A., Dekel, A., 2001, MNRAS submitted. [astro-ph/0111099]
Nusser A., da Costa L. N., Branchini E., Bernardi M., Alonso M., Wegner G., Willmer C., Pellegrin, P. 2000, MNRAS, 320, 21
Nusser, N., Dekel, A., Bertschinger, E., Blumenthal, G. R., 1991, ApJ, 379,6
Peebles P.J.E., 1980, The Large Scale Structure of the Universe, Princeton University Press, Princeton
Saunders W., Sutherland J., Maddox S., Keeble O., Oliver S., Rowan-Robinson M., McMahon R., Efstathiou G., Tadros H., White S., Frenk C., Carraminana A., Hawkins M. 2000, MNRAS, 317, 55
Seaborne M., Sutherland W., Tadros H., Efstathiou G., Frenk C., Keeble O., Maddox S., McMahon R., Oliver S., Rowan-Robinson M., Saunders W., White S. 1999, MNRAS, 309, 80
Sigad Y., Dekel A., Strauss M., Yahil A. 1998, ApJ, 495, 516
Sigad Y., Branchini E., Dekel A. 2000, ApJ, 540, 62
Somerville, R.S., Lemson, G., Sigad, Y., Dekel, A., Kauffmann, G., White, S.D.M., 2001, MNRAS, 320, 289
Strauss M., Willick J., 1995, Phys Rep., 261, 271
Teodoro L., Branchini E., Frenk C. 2000, in Cosmic Flows 1999: Towards an Understanding of Large-Scale Structures, eds. Courteau S., Strauss M., Willick J. (ASP Conference series) p. 242
Wiener, N. 1949, in *Extrapolation and Smoothing of Stationary Time Series*, (New York: Wiley)
Willick, J. A., Courteau, S., Faber, S. M., Burstein, D., Dekel, A., & Strauss, M. A. 1997a, ApJS, 109, 333
Willick, J., Strauss, M. 1998, ApJ, 486, 629
Willick, J., Strauss, M., Dekel, A., Kolatt, T. 1997b, ApJ, 486, 629
Zaroubi, S., Hoffman, Y., Dekel, A., 1999, ApJ, 520, 413
Zaroubi, S., Hoffman, Y., Fisher, K.B., & S. Lahav, O. 1995, ApJ, 449, 446
Zaroubi, S., Zehavi, I., Dekel, A., Hoffman, Y., & Kolatt T., 1997, ApJ, 486, 21
Zaroubi, S., Bernardi,M., da Costa, L.N., Hoffman, Y., Alonso, M.V., Wegner, G., Willmer, C.N.A., Pellegrini P.S., 2001, MNRAS, 326, 375.
Zaroubi S., 2002a, The proceedings of the XIII Recontres de Blois "Frontiers of the Universe", L. M. Celnikier, et al. (Editors). In press [astro-ph/0206052].
Zaroubi S., 2002b, MNRAS, 331, 901.