Online material

Appendix A. Details on data collection and analytical methods

A.1. Variables of the 'orchard design' group

The two variables *Density* and *DensityCV* were calculated from the measurement of the local planting density of all the N trees of the sample ($N \leq 100$ trees). The local planting density of the k^{th} tree (LD_k, ha^{-1}) was defined as follows:

$$LD_k = \frac{10^4}{d_1 d_2},$$

where d_1 (m) and d_2 (m) were the average distances of the tree to the two neighboring trees, respectively, in the row and between the rows.

Density was the mean of all the LD_k measurements. It indicates the overall planting density of the orchard. *DensityCV* was the coefficient of variation of all the LD_k measurements. It indicates the level of homogeneity of the orchard in terms of planting density (*i.e.*, a high value means a low homogeneity).

The 14 variables that describe the composition of the orchard in tree species (*i.e.*, *Mango, Orange, Grapefruit, Mandarin, Lemon, Papaya, Guava, OtherFruit*) and mango tree cultivars (*i.e.*, *Kent, Keitt, Boucodiékhal, Dieg bou gatt, Séwé and OtherMango*) were calculated based on the identification of the species of all the N trees of the sample ($N \leq 100$ trees) and the cultivars of mango trees. They were defined as follows:

$$i = \frac{N_i}{N},\ i \in \{Mango,\ Orange,\ Grapefruit,\ Mandarin,\ Lemon,\ Papaya,\ Guava,\ OtherFruit\},$$

$$j = \frac{N_{Mang,j}}{N_{Mango}},\ j \in \{Kent,\ Keitt,\ BDH,\ DBG,\ Séwé,\ OtherMango\},$$

where N_i is the number of trees of species i, N_{Mango} is the number of mango trees, and $N_{Mango,j}$ is the number of mango trees of cultivar j. All the tree species or mango cultivars present in less than six orchards were, respectively, gathered into the '*OtherFruit*' and '*OtherMango*' categories. In most of the cases, these species and cultivars represented less than 15% of, respectively, the trees and mango trees of the orchard.

A.2. Variables of the 'orchard vegetative state' group

The two variables *Cover* and *CoverCV* were calculated from the estimation of the local ground covering of all the N trees of the sample ($N \leq 100$ trees). The local ground covering of the k^{th} tree (LC_k, %) was defined as follows:

$$LC_k = \min\left(100; \frac{100\,\pi(d_f/2)^2}{d_1 d_2}\right),$$

where d_f (m) is the foliage diameter of the tree, and d_1 (m) and d_2 (m) are the average distances of the tree to the two neighboring trees, respectively, in the row and between the rows. The numerator value of the equation represents the area of the vertical projection of the foliage of the tree on the ground. The denominator value of the equation represents the area of ground available for the tree. *Cover* was the mean of all the LC_k measurements. It indicates the overall ground covering of the orchard by the tree canopy. Note that, because we fixed an upper bound of 100% for LC_k, the value of ground covering of the orchard by the tree canopy tends to be underestimated. *CoverCV* was the coefficient of variation of all the LC_k measurements. It indicates the level of homogeneity of the orchard in terms of ground covering (*i.e.*, a high value means a low homogeneity).

A.3. Variables of the 'hedgerow structure' group

During the survey, all the species of trees or shrubs present in a hedgerow were identified. A coefficient of abundance-dominance

Table AI.
Correspondence between the abundance-dominance coefficients (Braun-Blanquet scale) of a species *q* and the average ground covering of the species [32].

Abundance-dominance coefficient	Definition	Average ground covering (C_q, %)
5	Species covering > 75%, any abundance	87.5
4	Species covering from 50% to 75%, any abundance	62.5
3	Species covering from 25% to 50%, any abundance	37.5
2	Abundant species or species covering from 5% to 25%,	15
1	Little abundant species or species covering < 5%	3
+	Very little abundant species	0.5

taken from the Braun-Blanquet scale was assigned to each inventoried species and converted into an average ground covering percentage (*table AI*).

HedgeCover, the total hedgerow ground covering (%) by trees and shrubs, was the sum of the average ground coverings of all the inventoried species:

$$HedgeCover = \sum_{q=1}^{N} C_q ,$$

where C_q (%) is the average ground covering of the q^{th} species.

$Use_{k,\ k \in \{1,\ldots,5\}}$, the % of the total hedgerow ground covering by trees and shrubs of utility k, was the ratio between the sum of the average ground coverings of all the inventoried species of utility k and total hedgerow ground covering:

$$Use_k = 100 \sum_{q=1}^{N_k} C_{q,k} \Big/ \sum_{q=1}^{N} C_q ,$$

where N_k is the number of species of utility k and $C_{q,k}$ is the average ground covering of the q^{th} species of utility k.

4.1. A.4. Definition of the *v*-test criterion for qualitative variables

For a categorical variable and under the null hypothesis that the number of orchards for which the variable takes the category j is the same in cluster q (n_{jq}) and in the overall sample (n_j), the v-test is:

$$v - test = \frac{n_{jq} - n_q \dfrac{n_j}{N}}{\sqrt{n_q \dfrac{N - n_q}{N-1} \dfrac{n_j}{N} \left(1 - \dfrac{n_j}{N}\right)}} ,$$

where n_q denotes the number of orchards for the cluster q and N the total number of orchards. The higher the absolute value of the v-test, the more the category j of the categorical variable characterizes the orchards in cluster q. The sign of the v-test indicates if the category j is under- (v-test < 0) or over- (v-test > 0) represented among the orchards of the cluster q.

Appendix B. Complementary results

B.1. Definition and characterization of the orchard clusters

Figure B1 shows the results of the hierarchical clustering performed using the orchard design and management variables as active ones. The hierarchical clustering suggested a clustering into four orchard clusters.

The clusters were well separated by the first two principal components of the Hier-

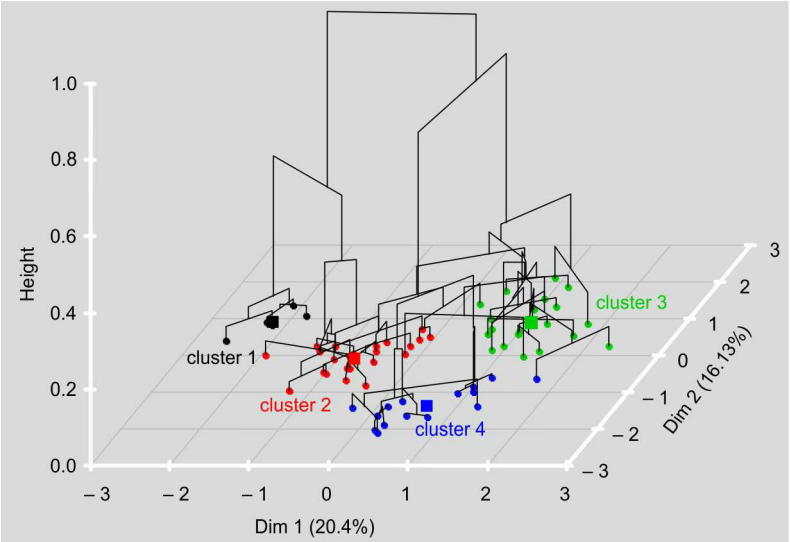

Figure B1.
Hierarchical clustering of the orchards obtained from the orchard planting design and management variables: 3-dimensional representation of the hierarchical tree and clusters on the map induced by the 1-2 Hierarchical Multiple Factor Analysis principal components.
Points represent individual orchards. Squares represent barycenters of individuals in each cluster.

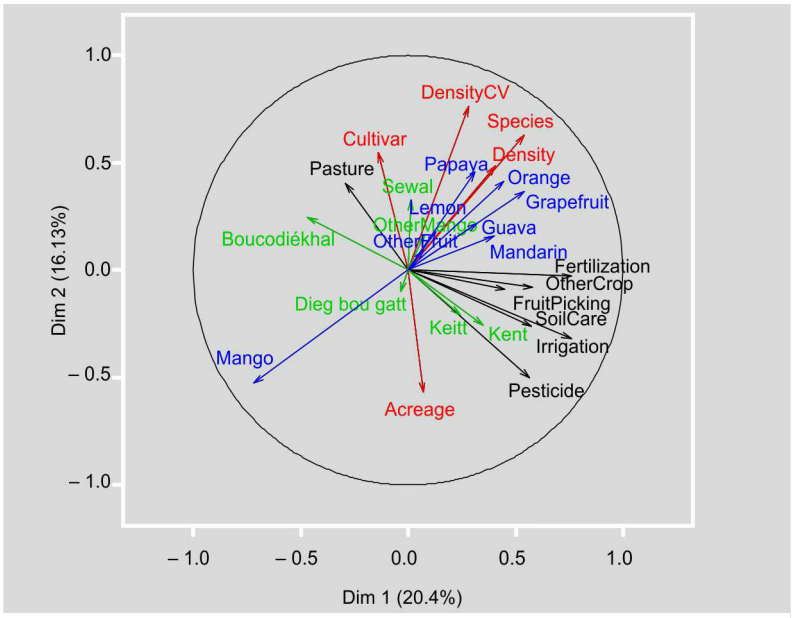

Figure B2.
Correlation plot of the orchard design and management variables with respect to the 1-2 Hierarchical Multiple Factor Analysis dimensions.
Black: variables of the management group; colored: variables of the planting design group structured into seven sub-groups, *i.e.*, the five sub-groups composed of a single variable (red), the sub-group representing the composition of the orchard in mango tree cultivars (green) and the sub-group representing the composition of the orchard in fruit tree species (blue).

archical Multiple Factor Analysis performed on the orchard design and management variables prior to clustering. Information on the contribution of the 'active' variables to the Hierarchical Multiple Factor Analysis components is given in *figure B2*.

A characterization of the orchard clusters with active and illustrative variables is given in *figure 2*.

The localization of the 64 orchards in the four geographic areas is given for each of the orchard clusters (or cropping systems) in *table BI*. The five orchards in cluster 1 were all located in Pékou and Pout. The orchards in cluster 2 were fairly equally located in the four geographic areas. The orchards in cluster 3 were mostly located in Sébikotane (85% of the orchards in cluster 3 *vs.* 34.4% of the orchards in the entire sample; $P < 1e{-}04$). The orchards in cluster 4 were mostly located in Notto (93.3% *vs.* 31.3%; $P < 1e{-}04$).

B.2. Definition and characterization of the hedgerow clusters

A typology of the hedgerows surveyed in 54 orchards was built based on the vegetative structure (abundance and composition in species' utilities) of their perennial part (*table Ie*). *Figure B3* shows the obtained partition of the orchards into four 'hedgerow' clusters. Only the hedgerow located in orchard n°35 changed from cluster 4 (*figure B3-A*) to cluster 2 after consolidation (*figure B3-B*). The consolidated clusters 1 to 4 contained, respectively, 4, 8, 17 and 10 orchards. These clusters were well separated on the 1-2 Hierarchical Multiple Factor Analysis principal components (*figure B3-B*). Information on the contribution of the hedgerow structure variables to these components is available in *figure 1-D*. A description of the clusters with active hedgerow variables is given in *table BII*.

Table BI.
Geographic localization of the orchards according to the orchard clusters (Senegal).

Orchards [(a)]	Geographic areas [(b)]			
	Notto (n = 20)	Pékou (n = 10)	Pout (n = 12)	Sébikotane (n = 22)
Entire sample (n = 64)	31.3	15.6	18.7	34.4
Cluster 1 (n_q = 5)	0	40	60	0
Cluster 2 (n_q = 24)	20.8	33.3 **	29.2	16.7 *
Cluster 3 (n_q = 20)	5 **	0*	10	85 ****
Cluster 4 (n_q = 15)	93.3 ****	0	0	6.7 *

[(a)] Orchard clusters are defined in section 3.2 (1: 'No-input mango diversified orchards', 2: 'Low-input mango orchards', 3: 'Medium-input citrus-predominant orchards', 4: 'Medium-input large mango- or citrus-predominant orchards').
[(b)] Values stand for the % orchards in the entire orchard sample or in an orchard cluster that were located in a given geographic area (*e.g.*, 20.8% of the orchards in the orchard cluster 2 were located in Notto). n is the number of orchards in the entire sample or sampled in a geographic area. n_q is the number of orchards belonging to a cluster. Asterisks indicate the p-value of the v-test (**** $P <$ 1e–04, *** $P <$ 1e–03, ** $P <$ 1e–02 and * $P <$ 5e–02) used to identify the variables that most characterize the orchard clusters.

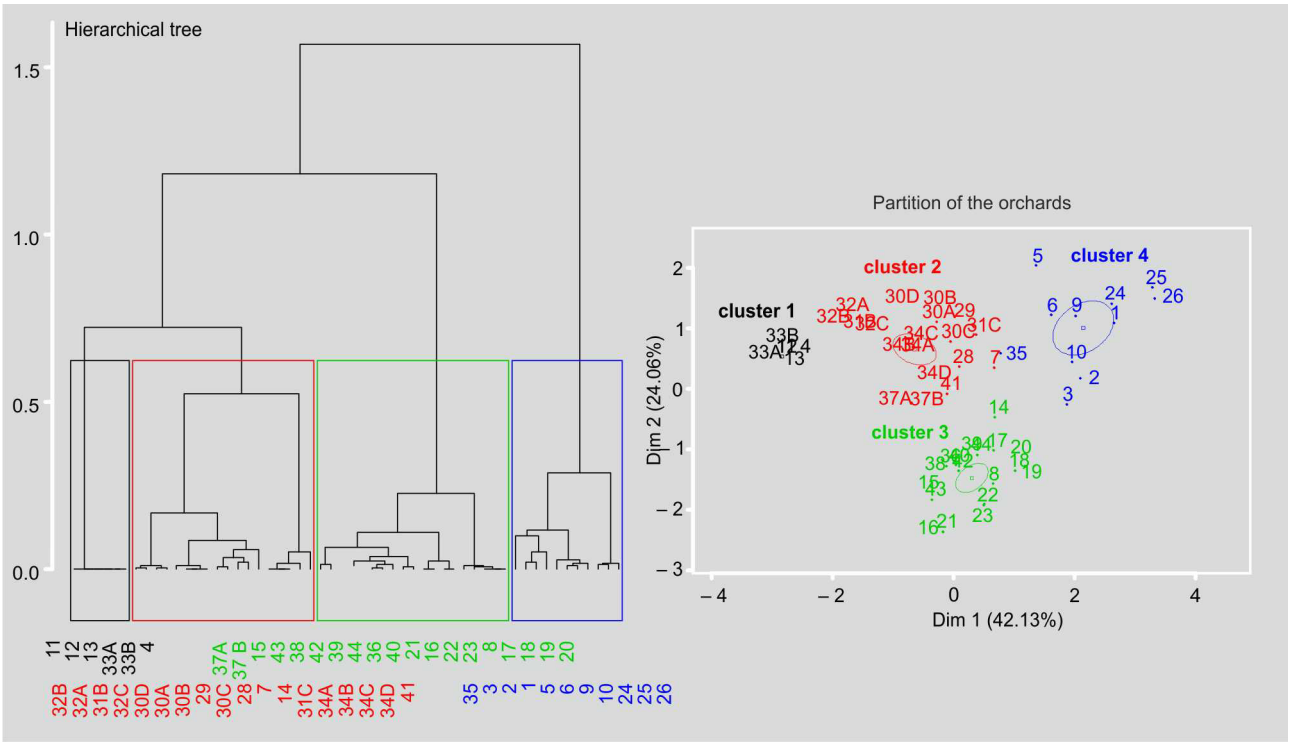

Figure B3.
Hierarchical clustering of the orchards obtained from the hedgerow structure variables: (A) hierarchical tree and (B) partition of the orchards, after consolidation, displayed on the factor map produced by the 1-2 Hierarchical Multiple Factor Analysis principal components.
Numbers stand for the identity number of the orchards in which the hedgerows were located. (A) Colored boxes enclose orchards of the same 'hedgerow' cluster. (B) Squares represent the barycenters of individuals in each cluster surrounded by their confidence ellipses at a 90% confidence level.

Table BII.
Characterization of the hedgerow clusters with active variables: mean, standard deviation (SD), minimal (Min) and maximal (Max); *n* stands for the number of orchards in the sample.

Variable	Overall hedgerows (*n* = 54)		Hedgerow cluster 1	Hedgerow cluster 2	Hedgerow cluster 3	Hedgerow cluster 4
	Mean ± SD	[Min-Max]	Mean ± SD			
Hedge	72.3 ± 33.4	[0–100]	0 ± 0 ****	57.9 ± 22.1 *	98.6 ± 5.9 ****	86.4 ± 20.5
HedgeSpecies	6.22 ± 3.72	[0–15]	0 ± 0 ****	5.16 ± 1.64	4.50 ± 1.62	11.6 ± 2.25 ****
HedgeCover	71.3 ± 41.6	[0–156]	0 ± 0 ****	60.6 ± 24.7	84.9 ± 36.3	106.4 ± 29.6 **
Use1	39.3 ± 33.4	[0–100]	Not available	12.7 ± 16.5 ****	75.3 ± 16.9 ****	26.5 ± 20.6
Use2	38.0 ± 32.6	[0–95]	Not available	64.7 ± 29.6 ****	11.6 ± 13.0 ****	34.9 ± 22.5
Use3	17.0 ± 15.5	[0–71]	Not available	21.2 ± 20.9	11.9 ± 9.6	18.2 ± 9.9
Use4	2.4 ± 4.4	[0–14]	Not available	0.58 ± 2.1 *	0.11 ± 0.47 **	9.2 ± 4.2 ****
Use5	3.5 ± 6.0	[0–25]	Not available	1.1 ± 1.8 *	1.0 ± 2.4 *	11.6 ± 7.6 ****

Asterisks indicate the *p-value* corresponding to the *v*-test (**** $P < 1e{-}04$, *** $P < 1e{-}03$, ** $P < 1e{-}02$ and * $P < 5e{-}02$) and identify the variables that most characterize the hedgerow clusters.