E

联合国粮食及农业组织 | Food and Agriculture Organization of the United Nations | Organisation des Nations Unies pour l'alimentation et l'agriculture | Продовольственная и сельскохозяйственная организация Объединенных Наций | Organización de las Naciones Unidas para la Alimentación y la Agricultura | منظمة الأغذية والزراعة للأمم المتحدة

COMMISSION ON GENETIC RESOURCES FOR FOOD AND AGRICULTURE

HIGHER-ORDER COMPOSITE INDICES FOR PLANT GENETIC RESOURCES FOR FOOD AND AGRICULTURE TARGETS

by
Francesco Caracciolo[1], Carlo Cafiero[2] and Stefano Diulgheroff[3]

This document has been prepared at the request of the Secretariat of the FAO Commission on Genetic Resources for Food and Agriculture, and in close collaboration with the FAO Plant Production and Protection Division, to facilitate the Commission's deliberations when it will review key issues on targets and indicators for plant genetic resources for food and agriculture at its Fifteenth Regular Session.

The content of this document is entirely the responsibility of the authors, and does not necessarily represent the views of the FAO or its Members.

[1] Department of Agriculture, University of Naples Federico II
[2] Statistics Division, FAO
[3] Plant Production and Protection Division, FAO

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Objective of the study

According to the Second Global Plan of Action for Plant Genetic Resources for Food and Agriculture (Second GPA), overall progress on its implementation will be monitored and guided by governments and other FAO members through the Commission on Genetic Resources for Food and Agriculture (the Commission). The Commission at its last session adopted targets and indicators for monitoring the implementation of the Second GPA and requested FAO to finalize the Reporting Format for monitoring the implementation of the Second (Reporting Format[4]) accordingly. It also requested FAO to elaborate higher-order composite indices (HCIs) for each of the plant genetic resources targets, basing them on data collected from the adopted indicators[5].

In response to the Commission's request, this document proposes three HCIs for the plant genetic resources targets. It outlines the steps required for constructing and using the three HCIs, highlighting their assumptions and limitations, and exploring their applicability at national, regional and global levels. Finally, it provides guidance with regard to the further refinement and optimization of the methodology.

This document is the result of a thorough review and development process that has involved a systematic review of the relevant literature and consultation with experts. It will provide guidance and recommendation on how to proceed on the construction of the HCIs, contributing to a better understanding of the technical complexity behind their development. The proposed approach aims to ensure consistency in the data collected to monitor trends over time of the Second GPA implementation, enabling the comparison of performance across countries and regional areas. Since HCI development necessarily involves steps where arbitrary or subjective decisions have to be made, one of the aims of the present document is to drive choices between different strategies in dealing with HCI development.

To conclude, this document cannot be considered exhaustive. It has to be considered as a starting point for further improvements in HCI methodology.

1.2. Development of a composite index in general

According to the OECD Handbook on Constructing Composite Indicators[6], "composite indicators are much like mathematical or computational models. As such, their construction owes more to the craftsmanship of the modeller than to universally accepted scientific rules for encoding." In the same terms of computational models, justification and final acceptance of a composite indicator or index "relies on negotiation and peer acceptance (Saltelli, 2007)[7]" as well as its suitability to the proposed use more than its scientific and methodological rigour and sophistication (Cherchye, 2007)[8]. Peer acceptance as well as their fitness for the intended purpose are therefore essential for composite indicators or indices.

In general terms, the development of a composite index follows an ideal order of pre-defined steps, including: a) the development of a theoretical framework, b) data selection, c) imputation of missing data, d) normalization, e) weighting and f) aggregation (Saisana and Saltelli, 2011)[9].

a) The theoretical framework consists of the theoretical background that provides the basis for selecting and combining variables into a composite index. It describes the multi-faced dimension to be measured and its relationship with the sub-components.

---

[4] CGRFA-15/15/Inf.9.

[5] CGRFA-14/13/Report, paragraph 27.

[6] Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2005). Handbook on constructing composite indicators: methodology and user guide (No. 2005/3), OECD publishing, Paris.

[7] Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research* 81(1), 65-77, page 70.

[8] Cherchye, L., et al. (2007). Creating composite indicators with DEA and robustness analysis: the case of the technology achievement index. *Journal of the Operational Research Society* 59(2), 239-251.

[9] Michaela Saisana and Andrea Saltelli (2011). Rankings and Ratings: Instructions for Use. *Hague Journal on the Rule of Law*, 3, pp 247-268.

b) Data selection is the process for the identification of the variables that allow the overall phenomenon addressed by the composite index to be captured. Variables should be selected on the basis of their specificity, measurability, availability, relevance and timeliness.

c) Imputation of missing data is the procedure to achieve the completeness of data required for computing the index.

d) Normalization is performed in order to render the variables comparable and aggregable, which can be expressed through different units of measure or scales.

e) Weighting is a judgment process that determines the contribution of each variable to the composite index. Weighting schemes might have significant effect on the overall composite index. The assignation of weights largely depends on views of the society and political standpoints. Most composite indicators rely on equal weighting (EW), i.e. all variables are given the same weight. Nevertheless, even the decision that all the variables are equally important in defining the composite index should be the outcome of a participatory method.

f) Finally, aggregation combines the weighted variables into one composite index. One of the most widespread aggregation procedures is the linear summation of weighted and normalized individual indicators.

During each of these steps, different choices are possible and the choice in one step may have important implications for the following steps. The choices depend on the aim and the specific characteristics of the indicators and together they define the overall modelling approach.

2. DEVELOPMENT OF HIGHER-ORDER COMPOSITE INDICES

2.1. Steps towards the three higher-order composite indices

In this paragraph details of the methodology for HCI computation will be discussed. In particular each step toward the development of HCI, highlighting pros and cons of the proposed model, will be specifically described. The three HCIs aim to aggregate ideally multi-faced concepts, corresponding to the 18 priority activities (PAs) of the Second GPA, into wider and primary dimensions, matching the three mutually supportive targets (*PGRFA Conservation*, *PGRFA Sustainable Use* and *PGRFA Institutional and Human Capacities*).

The proposed model explicitly takes into account the existence of a wide range of methodological approaches adopted by researchers, as well as the potential drawbacks of underlying indicators. Moreover, the model will be developed to fit the specific hierarchical or nested structure designed by the Second GPA for linking indicators to priorities and targets.

Even though HCIs are mathematical models and their development cannot be expressed without referring to any mathematical formulation and considering the statistical structure of whatever information source possible, a language without intensive use of mathematical notations and statistical background will be used.

The following notation will be adopted throughout: let $i_{n,c}^{t}$ be the value of the *n*-th indicator (with *n* =1,…,N) for the *c-th* country (with *c* = 1,…,C) at time t (with t = 1,…,T); $I_{n,c}^{t}$, the normalized value of the indicator; $P_{g,c}^{t}$, the partial score for the *g-th* priority activity (with g = 1,…,G) and w_g the associated weight for aggregating PAs. Finally let $HCI_{h,c}^{t}$ be the higher-order composite index value for the *h*-th target (with h=1,…, 3).

2.2. The theoretical framework of the three higher order composite indices

The Second GPA, its priority activities as well as the targets and indicators adopted by the Commission provide the theoretical framework for the HCIs: the Commission at its last session adopted 63 indicators to monitor the implementation of the 18 priority activities of the Second GPA and the following three mutually supportive targets10:

Target 1 - PGRFA Conservation. By 2020, an increasing proportion of the genetic diversity of cultivated plants and their wild relatives, as well as of wild food plant species, is maintained *in situ* , on farm and *ex situ* in a complementary manner;

Target 2 - PGRFA Sustainable Use: By 2020, there has been an increased use of plant genetic resources for food and agriculture to improve sustainable crop production intensification and livelihoods while reducing genetic vulnerability of crops and cropping systems;

Target 3 - PGRFA Institutional and Human Capacities: By 2020, many more people are aware of the value of plant genetic resources for food and agriculture and institutional and human capacities are strengthened to conserve and use them sustainably while minimizing genetic erosion and safeguarding their genetic diversity.

The purpose of HCIs is to assess progress towards the three PGRFA targets and to facilitate the comparison of performance across time, countries and regional areas. The implementation of the Second GPA as a whole contributes to the achievement of the adopted targets, and each priority activity covers a particular dimension of, and contributes to, one of the three targets (figure 1).

In particular, priority activities 1 to 7 of the Second GPA contribute to Target 1, priority activities 8-12 to Target 2, and priority activities 13-18 to Target 3. Progress in the implementation of each priority activity of the Second GPA is assessed through a set of indicators adopted by the Commission. Ideally, the indicators have to be first aggregated to give an overall score to the PA and then to the whole HCI.

[10] CGRFA-14/13/Report, Appendix C

HCI	**# Priority activities**	**# Indicators for priority activities**		
		Mean	**Min**	**Max**
PGRFA Conservation	7	3.4	3	5
PGRFA Sustainable Use	5	4.0	2	5
PGRFA Institutional and Human Capacities	6	3.2	2	5

Figure 1. Number of indicators for priority activities and number of priority activities for the three HCIs.

More specifically, the HCI for PGRFA Conservation will assess national progress on the implementation of seven priority activities related to surveying, inventorying (PA1) and collecting (PA5) of PGRFA, in addition to restoring crop systems after disaster situations (PA3) and the promoting of on-farm (PA4), in-situ (PA2) and ex-situ (PA6; PA7) conservation and management. The HCI for PGRFA Sustainable Use aims to monitor countries' priority activities for expanding characterization and evaluation of accessions (PA8), supporting plant breeding (PA9), promoting crop diversification (PA10) and the development and commercialization of new varieties (PA11) including seed production and distribution (PA12). Finally, the HCI for PGRFA Institutional and Human Capacities concerns national progress on strengthening PGRFA national programmes (PA13), networks (PA14) and information systems (PA15), developing monitoring systems for genetic diversity (PA16), strengthening human resource capacity (PA17), and raising public awareness on the importance of PGRFA (PA18).

The hierarchical or nested structure of HCIs is illustrated in figures 2 to 5.

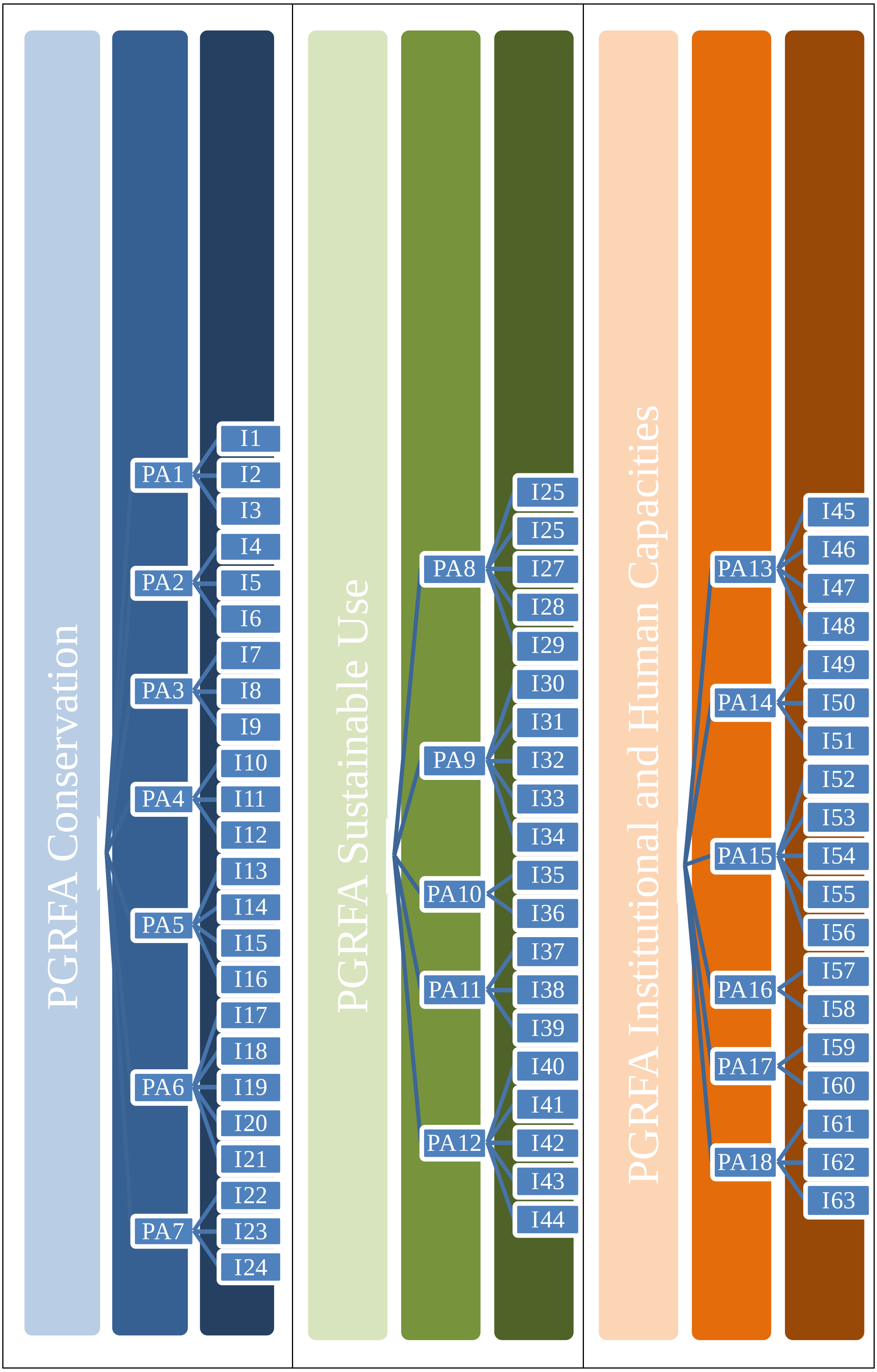

Figure 2. Hierarchical structure of the three HCIs

PGRFA Conservation Target	PA1. Surveying and inventorying plant genetic resources for food and agriculture	I1. Number of *in situ* (including on farm) surveys/inventories of PGRFA carried out I2. Number of PGRFA surveyed/inventoried I3. Percentage of PGRFA threatened out of those surveyed/inventoried
	PA2. Supporting on-farm management and improvement of plant genetic resources for food and agriculture	I4. Number of farming communities involved in on-farm PGRFA management and improvement activities I5. Percentage of cultivated land under farmers' varieties/landraces in areas of high diversity and/or risk I6. Number of farmers' varieties/landraces delivered from national or local gene banks to farmers (either directly or through intermediaries)
	PA3. Assisting farmers in disaster situations to restore crop systems	I7. Number of households that received seeds for planting as an aid after disaster situations I8. Percentage of seed produced at the local level out of that made available through disaster response interventions I9. Existence of disaster risk management policies for restoring crop systems that include seed security provisions
	PA4. Promoting *in situ* conservation and management of crop wild relatives and wild food plants	I10. Percentage of national *in situ* conservation sites with management plans addressing crop wild relatives and wild food plants I11. Number of crop wild relatives and wild food plants *in situ* conservation and management actions with institutional support I12. Number of crop wild relatives and wild food plant species actively conserved *in situ*
	PA5. Supporting targeted collecting of plant genetic resources for food and agriculture	I13. Existence of a strategy for identification of gaps in national gene bank holdings and for targeted collecting missions to fill identified gaps I14. Number of crops conserved in the national gene bank(s) that require targeted collecting I15. Number of targeted collecting missions in the country I16. Number of accessions resulting from targeted collecting missions in the country
	PA6. Sustaining and expanding *ex situ* conservation of germplasm	I17. Trend in annual capacity for sustaining *ex situ* collections I18. Number of crops conserved *ex situ* under medium or long-term conditions I19. Number of species conserved *ex situ* under medium or long-term conditions I20. Number of accessions conserved *ex situ* under medium or long-term conditions I21. Percentage of *ex situ* accessions safety duplicated
	PA7. Regenerating and multiplying *ex situ* accessions	I22. Percentage of *ex situ* accessions in need of regeneration for which a budget for regeneration does not exist I23. Number of *ex situ* accessions regenerated and/or multiplied I24. Percentage of *ex situ* accessions in need of regeneration

Figure 3. "PGRFA Conservation" - HCI description and specification

PGRFA Sustainable Use	PA8. Expanding the characterization, evaluation and further development of specific collection sub-sets to facilitate use	I25. Average number of morphological traits characterized per accession for the *ex situ* collections I26. Number of publications on germplasm evaluation and molecular characterization I27. Number of trait-specific collection subsets published I28. Number of accessions distributed by gene banks to users of germplasm I29. Number of samples distributed by gene banks to users of germplasm
	PA9. Supporting plant breeding, genetic enhancement and base-broadening efforts	I30. Number of crops with active public pre-breeding and breeding programmes I31. Number of crops with active private pre-breeding and breeding programmes I32. Number of breeding activities oriented to small scale farmers, villages or traditional communities I33. Number of active public crop breeders I34. Number of active private crop breeders
	PA10. Promoting diversification of crop production and broadening crop diversity for sustainable agriculture	I35. Number of programmes/projects/activities to increase genetic heterogeneity of crop species and diversity within the agro-ecosystem I36. Number of new crops and/or wild species introduced into cultivation
	PA11. Promoting development and commercialization of all varieties, primarily farmers' varieties/landraces and underutilized species	I37. Existence of national policies that promote development and commercialization of farmers' varieties/landraces and underutilized species I38. Number of programmes/projects/activities promoting development and commercialization of all varieties. I39. Number of farmers' varieties/landraces and underutilized species with potential for commercialization identified
	PA12. Supporting seed production and distribution	I40. Number of new varieties released I41. Number of formal/registered seed enterprises I42. The least number of varieties that together account for 80% of the total area for each of the five most widely cultivated crops I43. Percentage of area supplied with seed meeting the quality standard of the formal seed sector for the five most widely cultivated crops I44. Existence of a national seed policy and seed laws

Figure 4. "PGRFA Sustainable Use" - HCI description and specification

PGRFA Institutional and Human Capacities	Priority Activity	Indicators
	PA13. Building and strengthening national programmes	I45. Existence of a national entity (agency, committee, etc.) functioning as a coordination mechanism for PGRFA activities and/or strategies I46. Existence of a formally appointed national focal point or coordinator for PGRFA I47. Existence of a governmental policy framework and strategies for PGRFA conservation and use I48. Existence of a national information sharing mechanism for PGRFA
	PA14. Promoting and strengthening networks for plant genetic resources for food and agriculture	I49. Membership to a regional PGRFA network I50. Number of crop improvement networks in which national stakeholders are members I51. Number of publications produced by national stakeholders within the framework of networks
	PA15. Constructing and strengthening comprehensive information systems for plant genetic resources for food and agriculture	I52. Number of crop wild relatives conserved *in situ* and documented in a publicly available information system I53. Number of farmers' varieties/landraces cultivated on-farm and documented in a publicly available information system I54. Number of accessions from *ex situ* collections documented in a publicly available information system I55. Number of released varieties documented in a publicly available information system I56. Participation in publicly accessible, international/regional PGRFA information systems
	PA16. Developing and strengthening systems for monitoring and safeguarding genetic diversity and minimizing genetic erosion of plant genetic resources for food and agriculture	I57. Existence of national systems to monitor and safeguard genetic diversity and minimize genetic erosion I58. Number of remedial actions resulting from the existing national systems to monitor and safeguard genetic diversity and minimize genetic erosion
	PA17. Building and strengthening human resource capacity	I59. Existence of post-graduate, graduate and secondary educational and training programmes with incorporated aspects on PGRFA conservation and sustainable use I60. Percentage of staff whose skills in conserving and using PGRFA have been upgraded
	PA18. Promoting and strengthening public awareness of the importance of plant genetic resources for food and agriculture	I61. Existence of a public awareness programme promoting PGRFA conservation and utilization I62. Number of stakeholder groups participating in the implementation of the public awareness programme I63. Number of types of products developed to raise public awareness

Figure 5. “PGRFA Institutional and Human Capacities” - HCI description and specification

2.3. Data selection

Following the recommendation of the Commission after the adoption of the Second GPA, FAO, including the Secretariats of the Commission and the International Treaty, in collaboration with the Global Crop Diversity Trust and the CGIAR, undertook a revision of the 83 indicators for monitoring the implementation of the first GPA, in the light of the change introduced in the Second GPA and taking into consideration, in particular, the availability and accessibility of data required as well as the importance of maintaining continuity in reporting on the implementation of the GPA through a country-led participatory process[11]. The resulting draft set of indicators were subsequently revised by the ITWG-PGR at its Sixth Regular Session and the Commission at its Fourteenth Regular Session, which finally adopted them.

Although a lot of data on PGRFA have been collected by FAO most of the indicators adopted by the Commission are being used for the first time and the HCI model therefore has to be developed while no relevant data are currently available The overall quality of the HCIs in terms of accuracy and credibility depends greatly on the quality of basic data.

Analysis of data source quality is thus a necessary task for obtaining reliable HCIs and cannot be considered a one-off activity. Analytically, a good indicator has to be **SMART**: it "should be clearly and unambiguously defined (**S**pecific), be measurable in qualitative or quantitative terms (**M**easurable), be achievable in terms of the available resources (**A**chievable), be relevant to the issue in hand (**R**elevant) and be sensitive to changes within policy time-frames (**T**ime-bound) (Niemeijer and de Groot, 2008)[12]". It is very important to stress once again that a lack of quality in underlying indicators could limit the overall soundness and robustness of the obtained HCIs. For this reason, the underlying indicators will be described according to their compliance with the **SMART** rule. This procedure helps to identify sources of potential drawbacks, justifying specific methodological choices or suggesting procedures for further improvement of data collection.

As regards the property "*specific*", great attention may be paid to lessen potential bias coming from ambiguity/interpretation problems within the questions. Since "what is badly defined is likely to be badly measured[13]", the indicators should be based on shared definitions and concepts across countries and cultures. Likely, problems related to comprehension of the questions could be discussed by experts in specific focus groups, in order to ensure consistency across countries and thus cross-country comparability. If necessary, a detailed description including different ways to express key concepts could be provided. Furthermore, since some indicators are likely to be dependent on exogenous and environmental factors, it might be necessary to scale some indicators by an appropriate size measure such as population, GDP, land area or total accessions, to ensure an objective comparison across countries.

The "*measurable"* property regards the distance or "errors" between the data measured by indicators as collected through the Reporting Format, and the unknown real measure of the phenomenon. Indicators might suffer from a significant lack of available data – especially for baseline years. Countries may not provide the relevant data because they are not available or an indicator does not apply to them. The Reporting Format at the request of the Commission gives countries the option to skip reporting of individual indicators (e.g. inapplicable indicators and/or data not available).

Besides non-response, other sources of bias might include insufficient survey/sampling coverage in less developed areas. Once measured, the units of measures of indicators included in the Reporting Format obviously differ from each other, referring to very different domains of application. Furthermore, indicators largely differ from each other in the way they measure the domain under investigation (percentages, binary indicators, absolute values and indices/percentage change) (figure 6).

---

[11] CGRFA-14/13/4.1 Rev.1, paragraph 8

[12] Niemeijer, D., & de Groot, R. S. (2008). A conceptual framework for selecting environmental indicator sets. Ecological indicators, 8(1), 14-25.

[13] Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2005). Handbook on constructing composite indicators: methodology and user guide (No. 2005/3), OECD publishing, Paris. p. 12.

	Share/ percentage	Dichotomic/Binary	Positive integer (Natural number)	Indices - Percentage change	**Total**
PGRFA Conservation	16	2	5	1	24
PGRFA Sustainable Use	10	2	8		20
PGRFA Institutional and Human Capacities	4	7	8		19

Figure 6. Numerical nature of indicators

The Working Group, at its Sixth Session, noted that many indicators may not be easily *achievable* and recommended that the Reporting Format should allow respondents to show where specific indicators are not applicable.[14] The scope of the Second GPA is wide as its 18 PAs range from PGRFA conservation to use though capacity building. The need to limit the workload for countries and yet to collect a set of data that allow adequate assessment of the progress in their implementation has been taken into consideration throughout the identification and revision process of the indicators, during which National Focal Points (NFPs) and experts have been fully engaged. Nonetheless, some questions of the Reporting Format require complex answers and time-costly tasks to be properly addressed particularly if these are carried out by only one person per country. The participatory approach, involving different national stakeholders in data collection and reporting, together with the incorporation of existing data sources applying international standards, is indeed essential to reduce and distribute the workload, widen the coverage and limit the overall cost of data collecting and updating of the indicators over time.

The *relevance* characteristic refers to the quality of the indicators in representing/fitting the overall purpose and dimension of the HCI. As regards this property, one of the potential limitations of the adopted indicators is that their “optimum” target values are not always obvious or necessarily the same for all the countries. This uncertainty may involve even the correct understanding of the sign of the relationship between the indicators and the corresponding HCI: in these few cases, indicator interpretation might seem at first sight ambiguous, or at least poorly-defined. Therefore, when necessary, statistical approaches for identifying the sign of the relation could be employed after completion of the data collection process[15]. The following figure reports the expected sign of the relation between each indicator and the corresponding dimension measured by the HCI (figure 7).

All the above mentioned sources of data heterogeneity and inaccuracy are expressly considered during the development of the methodology for computing HCIs. To this end, in order to address the potential drawbacks of the 63 indicators, expectation on data availability, as well as several sources of data inaccuracy and the need for contextualization, the model proposes countries to provide an expert judgement on the level of achievement or implementation of the underlying dimension of each indicator. The expert judgement, it is proposed, would be provided by the National Focal Point. The data provided and the calculation of the corresponding indicators continue to be essential as the NFPs are guided by their values in their expert judgement. The expert judgement will allow interpretation and meaningful contextualization and codification of the quantitative measures coming from the indicators, as well as the data collected for calculating them. It will also allow mitigation of the effects caused by the heterogeneity of the values of the indicators due to the different national and environmental contexts. This may help to increase consistency and applicability of the three HCIs, as well as their comparability across time and among countries. However, the credibility and objectivity of NFP responses and judgments should be a

[14] CGRFA/WG-PGR-6/12/REPORT, paragraph 10-11.
[15] This point will be discussed in section 4.1

prerequisite since they are likely to influence the overall confidence that could be placed in the final HCIs.

Indicators	Relation	Indicators	Relation	Indicators	Relation
I1	Positive	I22	Negative	I43	Positive
I2	Positive	I23	Positive	I44	Positive
I3	Negative	I24	Negative	I45	Positive
I4	Positive	I25	Positive	I46	Positive
I5	Positive	I26	Positive	I47	Positive
I6	Positive	I27	Positive	I48	Positive
I7	Positive	I28	Positive	I49	Positive
I8	Positive	I29	Positive	I50	Positive
I9	Positive	I30	Positive	I51	Positive
I10	Positive	I31	Positive	I52	Positive
I11	Positive	I32	Positive	I53	Positive
I12	Positive	I33	Positive	I54	Positive
I13	Positive	I34	Positive	I55	Positive
I14	Negative	I35	Positive	I56	Positive
I15	Positive	I36	Positive	I57	Positive
I16	Positive	I37	Positive	I58	Positive
I17	Positive	I38	Positive	I59	Positive
I18	Positive	I39	Positive	I60	Positive
I19	Positive	I40	Positive	I61	Positive
I20	Positive	I41	Positive	I62	Positive
I21	Positive	I42	Positive	I63	Positive

Figure 7. Expected relation between each indicator and the overall HCI.

2.4. Imputation of missing data

In order to address the challenge derived from the potential unavailability of relevant data the proposed model minimizes the use of missing data substitution procedures. Since from two to five indicators are available for each PA, indicators with missing data are simply discarded. Imputation of missing data will be performed only if a country suffers from the absence of all the available indicators within a specific PA and on the indicator with the largest data availability.

Missing data are here considered as missing completely at random (MCAR[16]) because it is assumed that the reason for lack of data is unknown. The proposed model would impute missing data from the sample unconditional median (if continuous) or mode (if categorical) of the previous years' values for the given indicator for the country *c*. In the absence of previous information on $i_{n,c}^{t}$, the median (if continuous) or mode (if categorical) will be calculated from the sample of the indicator *n* recorded from the other countries in the same geographical subregion[17] at time *t*.

2.5. Data normalization

Having obtained information for at least one indicator for each PA for each country, the values of the indicators $i_{n,c}^{t}$ will be normalized (figure 8, phase a). Recorded values will be normalized on an ordered categorical scale from 1 to 9, 1 being the worst and 9 being the best. The use of a nine-category scale is driven by the desire to simplify their use and interpretation. A similar scale is already applied to measure

[16] Data are MCAR if the probability that data being missing is independent of the value of both the observed and unobserved data. Heitjan, D.F., and S. Basu. 1996. "Distinguishing "Missing at Random" and "Missing Completely at Random"." *The American Statistician* 50(3):207-213, p. 207.
[17] Regions and sub-region compositions are illustrated in figure 16.

progress in the implementation of the Global Plan of Action for Animal Genetic Resources[18]. An average of the normalized indicators of each PA is calculated to produce a PA score. For the first round, or for t = 1, the normalized score $I_{n,c}^{t}$ could be based either on:

A1- on a qualitative rule (expert assessment of the NFP). In this case, each country's indicators are scored through an anchored rating scale from 9 "*fully achieved*" or "*full degree of implementation*" to 1 "*not achieved*" or "*low degree of implementation*";

and as an alternative to subjective assessments,

B1- on a quantitative or performance rule (percentiles of the distribution of the indicator across countries for the year t[19]). Should $i_{n,c}^{t}$ assume dichotomic/binary values, zero values will be normalized to 1, while 1's to 9.

In presence of a negative sign of the relationship between $i_{n,c}^{t}$ and the overall HCI (paragraph 1.3; figure 7), the corresponding normalized indicator $I_{n,c}^{t}$, if obtained using a quantitative or performance rule (b), has to be "*reversed*": $I_{n,c}^{t}(\boldsymbol{R}) = 10 - I_{n,c}^{t}$. For example, with the reverse operation, what was scored as 1 becomes 9, while what was scored as 9 becomes 1.

This procedure allows benchmark values of the indicators to be generated in order to assess in the following rounds whether any changes occur in the country over time. For the following rounds, or for t >1, again two different data normalization procedures are suggested, depending on the choice made for the first round:

A2 - the same normalization scheme (**A1**) using a nine-category scale through a subjective assessment of the NFP or a panel of experts;

and as an alternative to the subjective assessments,

B2 - a normalization by using the proportion of annual differences over consecutive years, $\frac{i_{n,c}^{t}-i_{n,c}^{t-1}}{i_{n,c}^{t-1}}$, with $I_{n,c}^{t}$ being calculated as $I_{n,c}^{t}= I_{n,c}^{t-1} \times \left(1 + \frac{i_{n,c}^{t}-i_{n,c}^{t-1}}{i_{n,c}^{t-1}}\right)$.

With at least one normalized indicator $I_{n,c}^{t}$ obtained for each country and for PA, a simple average of the available normalized indicators for each country will provide the PA scores (figure 8, phase b) $P_{g,c}^{t} = \frac{1}{M_{p,c}}\sum_{n}^{M_{p,c}} I_{n,c}^{t}$, where $2 \leq M_{g,c} \leq 5$ corresponds to the number of the available normalized indicators for the *c-th* country and the g-*th* priority. PA scores are now ready to be weighted in order to calculate the corresponding higher-order composite index, $HCI_{h,c}^{t}$.

2.6. Weighting scheme and aggregation

At this stage, the next important step is to aggregate individual priority activities scores to obtain a final HCI score. The theoretical foundation of weighting and aggregation schemes was mainly explored by multi-criteria analysis (MCA) that formally studies empirical models in decision-making processes. Several aggregation methods have been proposed in the literature[20] and they can be used for this purpose. Among the alternatives, the linear and additive aggregation scheme is proposed here, following the

[18] CGRFA/WG-AnGR-7/12/7 paragraph 12

[19] The top 12.5% of the indicators distribution across countries score 9; 8 the indicators between the 25th percentile and the 12.5th percentile; 7 the indicators between the 37.5th and the 25th; 6 the indicators between the 50th and the 37.5th; 5 the indicators between the 62.5th and the 50th; 4 the indicators between the 75th and the 62.5th; 3 the indicators between the 87.5th and the 75th, 2 the indicators recording the worst values.

[20] An extensive literature review on the topic is given by Booysen, F. 2002. "An Overview and Evaluation of Composite Indices of Development." Social Indicators Research 59(2):115-151, pages 126-129.

suggestion of Booysen (2002): "Studies aimed at presenting a simple and informative view of general well-being or at informing officials regarding particular issues tend to opt for relatively simpler methods, thus allowing for indices to be easily comprehensible and readily calculable".

Additive linear aggregation is the most popular technique: for example, it has been utilized to develop the Environmental Sustainable Index[21]. Weighted linear aggregation consists in the simple weighted summation of the scores on PAs (figure 8, phase c):

$$HCI_{h,c}^{t} = \sum_{g=1}^{G} w_g P_{g,c}^{t}$$

with $\sum_{g=1}^{G} w_g$ = 1 and $0 \leq w_g \leq 1$, for all g = 1,...G, c = 1,..., C, and t = 1,…,T.

The linear additive aggregation method implies an implicit satisfaction of the *preference independence condition*, or the mutually preferential independence of the priority activities. This allows the existence of an additive function for aggregating the marginal contribution of each priority separately. This assumption may be regarded as not free from drawbacks, since it assumes that among the different priority activities there are no interactions, and a high value scored in an activity can completely offset a low value scored in another PA.

Weight w_g can be considered as a judgment value (Singh et al., 2007)[22] and it indicates the contribution of each PA to the overall HCI score. Thus, weighting schemes might have a significant effect on the overall HCI, impacting greatly on effective countries' assessment. The set of weights to assign to different priorities should largely depend on the views of society and political standpoints. Also the decision that all the priorities are equally important in defining the HCIs should be the outcome of a participatory method. Thus, weighting methods that include the viewpoints of different stakeholders, experts and politicians are to be preferred wherever possible.

Among the many participatory weighting methods, the analytic hierarchy process (AHP) could be followed (Saaty, 1988 [23]). AHP is a widely accepted method which controls for internal consistency for priority assessment. Based on a set of stated preferences of experts from rotating pair-wise comparison of priorities, this technique provides coherent and robust weights. Alternatively, a more simple budget allocation approach (BAL) could be followed: after stimulating the experts with the question "On what basis shall it be decided to allocate *x* dollars to activity A instead of activity B" (Key, 1940)[24], the weights are calculated by the experts' allocation of a fixed budget over the priorities.

In the absence of consensus, or in the presence of technical difficulties or time constraint for adopting the above weighting schemes, equal weighting (EW) can be employed, i.e. all the priority activities are given the same weight. However, EW adoption in the quantitative or performance approach (B1) does not necessarily imply that all the priority activities have the same "importance". Empirically, some priority activities could be *over-weighted* according to the statistical nature of the data source (e.g. presence of correlation between the priority activity scores) (Paruolo *et al*., 2013).

[21] The index is based upon a set of 68 basic indicators, comprised by 21 core indicators. The environmental sustainability index score for each country is the mean value for the 21 factors.
[22] Singh, R. K., Murty, H. R., Gupta, S. K., & Dikshit, A. K. (2007). Development of composite sustainability performance index for steel industry. *Ecological Indicators*, 7(3), 565-588.
[23] Saaty, T. L. (1988). What is the analytic hierarchy process?. Springer Berlin Heidelberg.
[24] Key, V. O (1940) ‘The lack of a budget theory’ American Political Science Review 34 (6), 1137-1144.

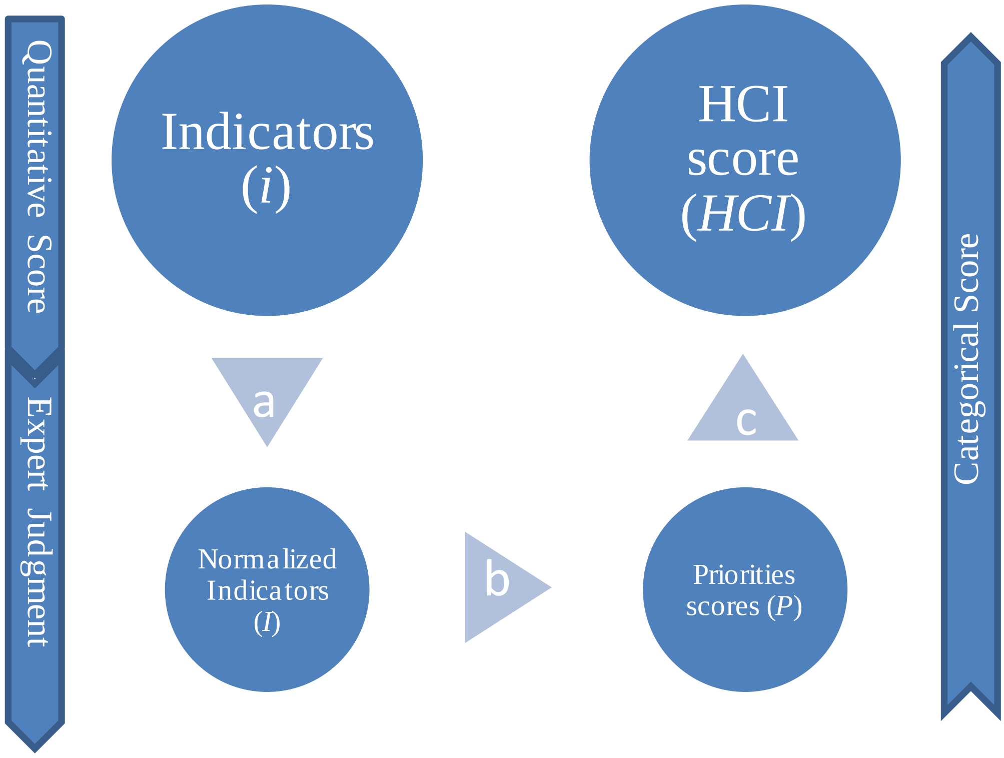

Figure 8. Synthetic scheme for HCI development

2.7. The proposed model in synthesis

In synthesis, four objectives have led to the proposed approach: (i) to minimize the use of missing data substitution procedures;(ii) to allow comparison of indicators (normalization) across countries by means of ordered categorical scales; (iii) to aggregate normalized indicators into PAs by computing averages, and then into HCI by performing weighted linear aggregation (aggregation), preferably according to a participatory method that could incorporate expert or stakeholder opinion (weighting); (iv) to provide an easily comprehensible and readily calculable method for HCI computation and graphical visualization, ensuring comparability of the HCIs over time and across countries.

In conclusion, the path of HCI development follows the process as schematized in figure 8.

3. USE OF THE COMPOSITE INDICES

3.1. Benchmarking and cross-country analysis

As illustrated in the introductory chapter, the primary use of the three HCIs is to provide a simplified and concise assessment, at country, regional and global levels, of the progress towards the three PGRFA targets, and of the implementation of the Second GPA and its PAs. Thus, once calculated, HCIs should clearly show to stakeholders, practitioners and policy-makers the degree of implementation of the state of the implementation of the various priorities of the second GPA at country/regional level. Adoption of the normalization procedure suggested in the previous chapter was indeed driven by the willingness to simplify use and interpretation of the HCIs. Furthermore, besides the HCI scores, a single score for each priority activity and country on an anchored scale ranging from '1 –Low implementation', to '9 - High implementation' will be generated.

The scores calculated at t=1 will provide the benchmark data against which future GPA implementation changes can be compared at all the levels where scores are calculated (indicators, priority activities, higher-order composite indexes). In detail, $I_{n,c}^{t=1}$ will provide the benchmark level for the n-th indicator for the c-th country, while the average of the available $I_{n,c}^{t=1}$ within each priority,

$\frac{1}{M_{p,c}}\sum_{n}^{M_{p,c}} I_{n,c}^{t=1}$, will provide the benchmark score of the *g*-th priority activity $P_{g,c}^{t=1}$. Finally, $HCI_{h,c}^{t=1} = \sum_{g=1}^{G} w_g P_{g,c}^{t=1}$ is the corresponding benchmark level of the *h*-th higher order composite index.

The approach for presenting HCIs is no minor issue. For example, it is possible to tabulate the HCI scores for each country (e.g. eight countries called A, B..., H.) as a table of values in descending ranking order, or following an alphabetical and regional order[25] (figure 9). Moreover, bar charts can be used to compare HCI scores between countries (figure 10).

	PA1	PA2	PA3	PA4	PA5	PA6	PA7	**HCI1**	PA8	PA9	PA10	PA11	PA12	**HCI2**	PA13	PA14	PA15	PA16	PA17	PA18	**HCI3**
Country A	2	5	2	7	2	8	5	**4.4**	1	9	1	8	3	**4.4**	3	8	2	5	1	1	**3.3**
Country B	6	7	5	4	1	2	7	**4.6**	2	7	7	2	2	**4.0**	3	6	1	6	5	3	**4.0**
Country C	5	3	3	4	5	3	1	**3.4**	3	2	4	3	4	**3.2**	7	2	6	4	6	4	**4.8**
Country D	5	1	8	8	3	2	7	**4.9**	6	6	5	2	5	**4.8**	9	6	5	2	7	8	**6.2**
Country E	6	4	3	1	7	7	3	**4.4**	8	6	3	5	2	**4.8**	7	3	7	5	5	3	**5.0**
Country F	1	9	1	6	1	3	4	**3.6**	4	5	3	1	7	**4.0**	6	7	1	9	3	5	**5.2**
Country G	5	5	4	6	5	4	2	**4.4**	2	1	4	5	6	**3.6**	4	5	4	7	2	6	**4.7**
Country H	1	7	6	6	1	8	2	**4.4**	8	4	2	6	3	**4.6**	1	6	7	6	5	5	**5.0**

Figure 9. Example of presentation of PA and higher order composite index scores.

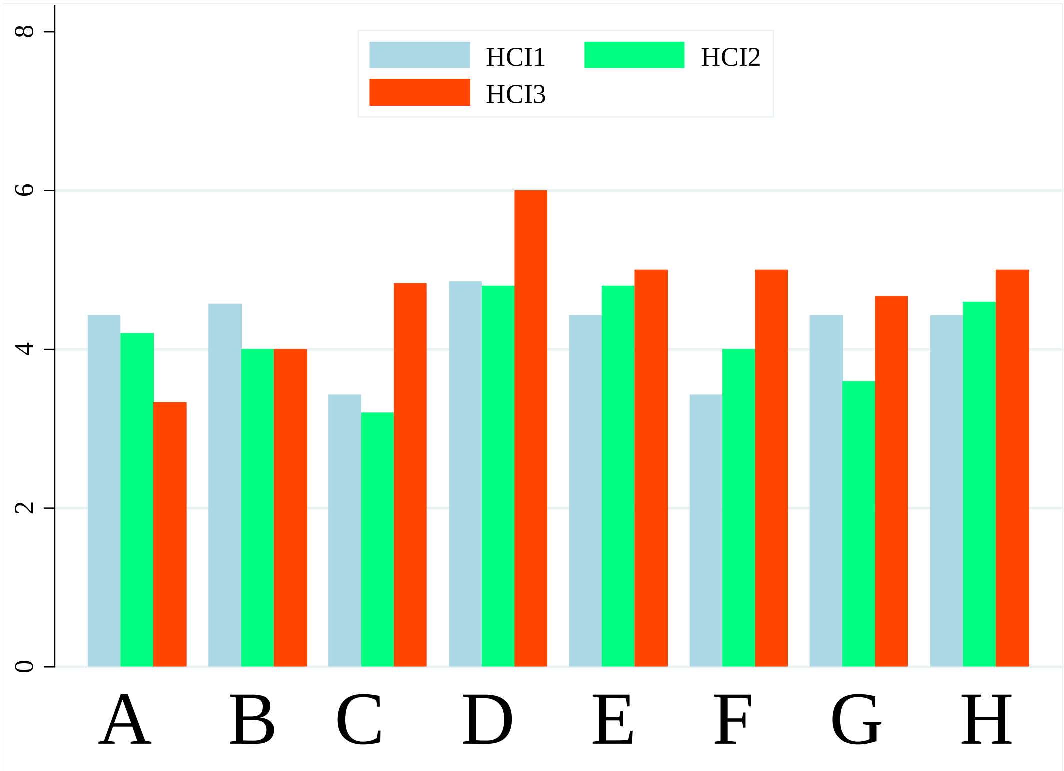

Figure 10. Example of bar chart presentation of HCIs

[25] See for example CGRFA/WG-AnGR-7/12/Inf.3, Annex 3, p. 89.

Alternatively, following the colour scheme based on traffic-lights developed to measure progress made in the implementation of the Global Plan of Action for Animal Genetic Resources[26], the scores can be categorized into eight ordered categories, represented by eight colours – three shades of red (indicating a low degree of implementation), two of yellow (indicating medium levels of implementation) and three of green (indicating a full degree of implementation). Scores with the corresponding categories and colours are shown in figure 11, while figure 12 shows an example of scores presented by using colour categories.

Scores -Priorities and HCI	Indicator colour	Indicator level
1.00 - 1.99		Low
2.00 - 2.99		Low
3.00 - 3.99		Low
4.00 - 4.99		Medium
5.00 - 5.99		Medium
6.00 - 6.99		High
7.00 - 7.99		High
8.00 - 9.00		High

Figure 11. Scores, categories and colours for presenting the degree of implementation of the Second GPA**.**

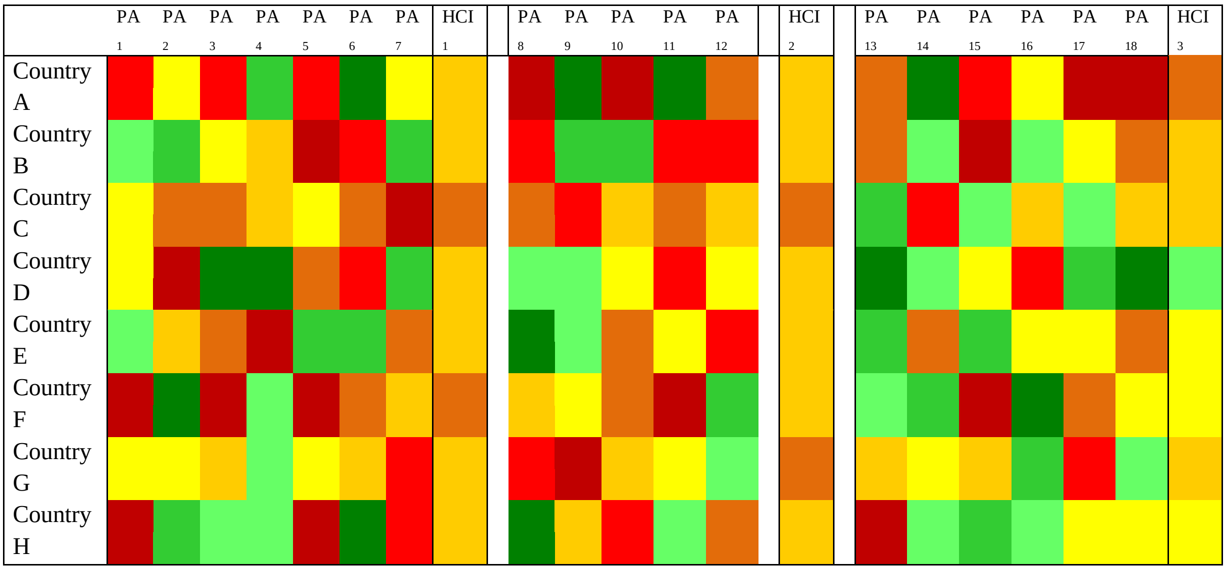

Figure 12. Presentation of priority activities and HCI scores.

3.2. Monitoring country performance over time

Having obtained the benchmark scores $P_{g,c}^{t=1}$ and $HCI_{h,c}^{t=1}$, the implementation of the second GPA of the *c*-th country can be easily measured and compared for any *t*-th period of time by computing $P_{g,c}^{t}$ and $HCI_{h,c}^{t}$, adopting whatever of the two normalization schemes we suggested in the previous section

[26] CGRFA/WG-AnGR-7/12/7 paragraph 12

(**A2** expert categorical scale or **B2** proportion of annual differences over consecutive years). Obviously, changes over time in priority activities and HCI scores for the *c*-th country can be calculated only when the scores are available for at least two different time points (figure 13).

Several presentations can be used to track country performance over the years using both absolute levels of scores or the percentage growth rates between the available years (figure 14). Furthermore, for each strategic target, the total number of priority activities indicating an increase can be counted, showing the country progress in terms of relative frequencies (figure 15a and 15b).

	PA_1	PA_2	PA_3	PA_4	PA_5	PA_6	PA_7	PA_8	PA_9	PA_{10}	PA_{11}	PA_{12}	PA_{13}	PA_{14}	PA_{15}	PA_{16}	PA_{17}	PA_{18}
Country A	2	5	2	7	2	8	5	1	9	1	8	3	3	8	2	5	1	1
Country B	6	7	5	4	1	2	7	2	7	7	2	2	3	6	1	6	5	3
Country C	5	3	3	4	5	3	1	3	2	4	3	4	7	2	6	4	6	4
Country D	5	1	8	8	3	2	7	6	6	5	2	5	9	6	5	2	7	8
Country E	6	4	3	1	7	7	3	8	6	3	5	2	7	3	7	5	5	3
Country F	1	9	1	6	1	3	4	4	5	3	1	7	6	7	1	9	3	5
Country G	5	5	4	6	5	4	2	2	1	4	5	6	4	5	4	7	2	6
Country H	1	7	6	6	1	8	2	8	4	2	6	3	1	6	7	6	5	5

time =1

	PA_1	PA_2	PA_3	PA_4	PA_5	PA_6	PA_7	PA_8	PA_9	PA_{10}	PA_{11}	PA_{12}	PA_{13}	PA_{14}	PA_{15}	PA_{16}	PA_{17}	PA_{18}
Country A	5	5	1	7	4	6	8	5	5	4	3	7	9	7	7	6	8	9
Country B	4	1	8	6	3	5	4	2	5	2	2	6	1	3	1	8	3	3
Country C	8	6	5	6	7	8	4	5	4	6	7	8	3	3	8	1	1	4
Country D	8	1	2	4	1	2	8	8	7	4	3	1	4	1	6	6	2	8
Country E	8	4	2	2	2	3	7	3	1	7	4	8	5	4	4	7	5	5
Country F	9	1	7	1	1	9	7	4	8	7	2	5	3	3	3	1	4	3
Country G	2	1	1	7	2	7	5	1	8	6	6	7	5	2	5	8	6	1
Country H	3	7	2	5	6	4	7	4	8	7	7	1	1	8	1	3	8	3

time =2

Figure 13. Example of a tabular presentation for assessing performance over time of priority activities and higher order composite index scores.

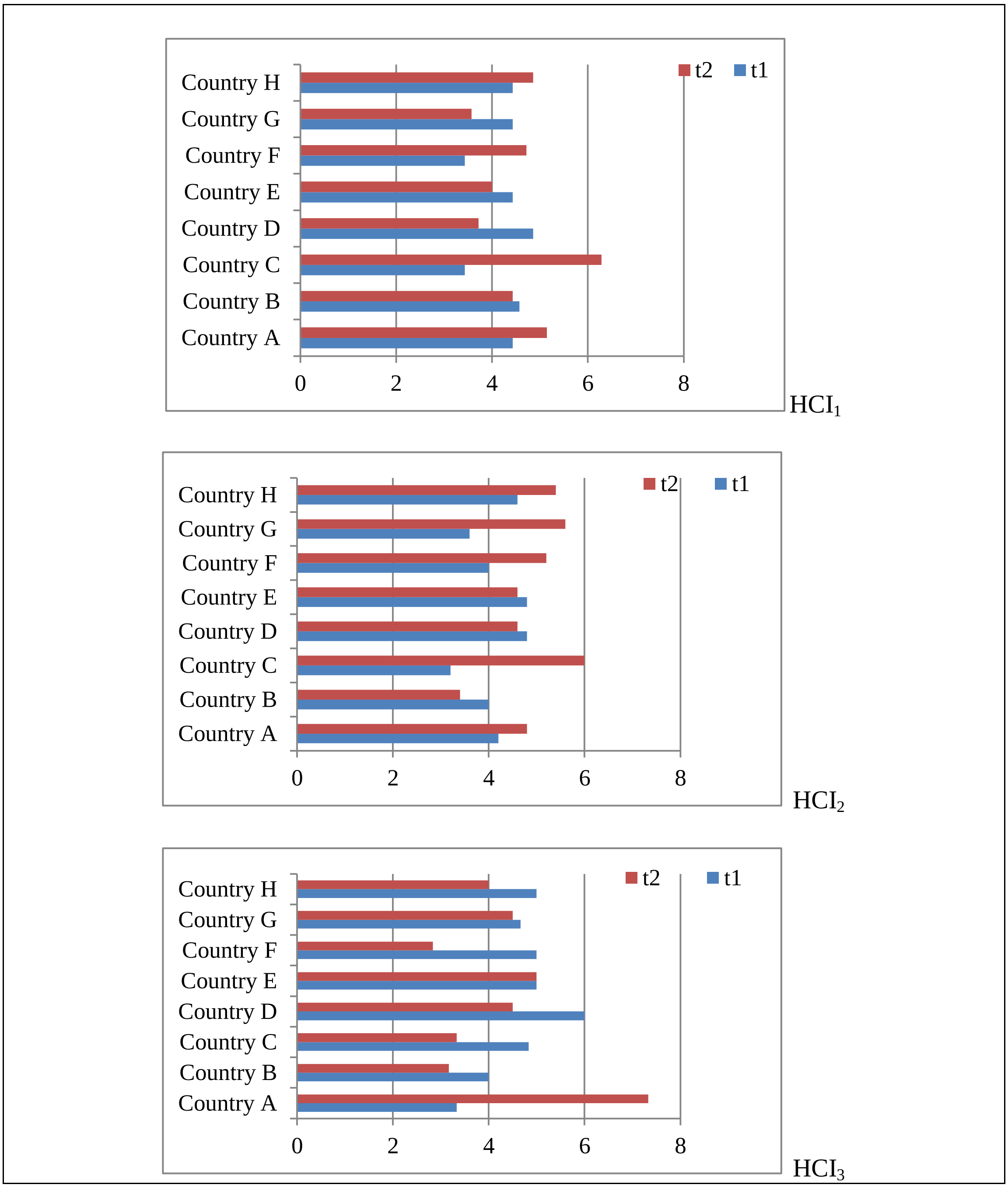

Figure 14. Example of a bar presentation of HCI change over time

	HCI_1	HCI_2	HCI_3
Country A	42.9	61.1	83.3
Country B	57.1	33.3	16.7
Country C	100.0	77.8	33.3
Country D	28.6	38.9	33.3
Country E	42.9	44.4	50.0
Country F	57.1	50.0	33.3
Country G	42.9	61.1	66.7
Country H	42.9	44.4	33.3

Figure 15a. Example of a tabular presentation of country progress in terms of relative frequencies

	HCI_1	HCI_2	HCI_3
Country A	=	✓	✓
Country B	=	✗	✗
Country C	✓	✓	✗
Country D	✗	✗	✗
Country E	=	=	=
Country F	=	=	✗
Country G	=	✓	✓
Country H	=	=	✗

Figure 15b. Example of traffic lights assessment of country progress

3.3. Aggregation at regional and global level

After having calculated the scores concerning priority activities ($P_{g,c}^{t}$) and higher-order composite indices ($HCI_{h,c}^{t}$) for individual countries, it is straightforward to monitor the state of implementation of the Second GPA at subregional, regional and global level. Any analysis comparing scores among regional levels must take into account the somewhat arbitrary nature of the definitions of the regions. The aggregation of countries within regions should be consistent, allowing policy-relevant comparisons that would be useful to public and private stakeholders. As regards this point, the same geographical classification adopted in the Statistics Division of the United Nations could be employed[27]. This specific aggregation is based on the M49 coding standard, and has been used to report advances towards achieving numerous millennial development goals worldwide. Following the M49 standard, countries are organized into 22 geographical sub-regions within five macro geographical regions (figure 16).

[27] See http://millenniumindicators.un.org/unsd/methods/m49/m49regin.htm for the complete organization of countries within macro geographical regions and geographical sub-regions.

Regions	Subregions
Africa	Eastern Africa
	Middle Africa
	Northern Africa
	Southern Africa
	Western Africa
Americas	Caribbean
	Central America
	South America
	Northern America
Asia	Central Asia
	Eastern Asia
	Southern Asia
	South-Eastern Asia
	Western Asia
Europe	Eastern Europe
	Northern Europe
	Southern Europe
	Western Europe
Oceania	Australia and New Zealand
	Melanesia
	Micronesia
	Polynesia

Figure 16. Regional and sub-regional composition

The most common methodology to aggregate country priorities and higher-order composite indices is to compute the simple average of the scores across countries:

$$\overline{P_g^t} = \frac{1}{S}\sum_{c=1}^{S} P_{g,c}^t; \qquad \overline{HCI_h^t} = \frac{1}{S}\sum_{c=1}^{S} HCI_{h,c}^t.$$

In alternative, weighted averages may be calculated using somewhat arbitrary weights (k_c) such as country's total area or country's population: $\overline{P_g^t} = \frac{\sum_{c=1}^{S} k_c \cdot P_{g,c}^t}{\sum_{c=1}^{S} k_c}$; $\overline{HCI_h^t} = \frac{\sum_{c=1}^{S} k_c \cdot HCI_{h,c}^t}{\sum_{c=1}^{S} k_c}$.

4. FURTHER REVISION OPTIONS

4.1. Inclusion or exclusion of indicators

Having developed HCIs, an assessment of the overall quality (appropriateness) of the collected indicators for being used in developing the HCIs could motivate further refinement and optimization of the methodology. This judgment can be supported either by qualitative (experts opinion) or quantitative (correlation structure of the data) analysis. Outcome of the assessment could recommend the exclusion of indicators or priority within the HCI computation (i.e. due to the high frequency of missing values), or provide options for further revision of the methodology.

Qualitative and visual inspection of the collected data, for example, may show that all the indicators within a priority activity were considered "*not applicable*" by an NFP for a specific country; moreover, the whole priority activity could be considered not applicable by the NFP. While this is a special case, it has to be considered explicitly for developing the HCIs consistently.

Let us consider priority $P_{1,c}^{t}$ not applicable by the NFP. In this case, the HCI score can be calculated by multiplying the score obtained without the priority times the quantity ($1+w_1$).

$$HCI_{h,c}^{t} = \left[\sum_{g=2}^{G} w_g P_{g,c}^{t}\right](1 + w_1)$$

Quantitative assessment of the quality of indicators is based on the statistical properties of the collected indicators. Multivariate analysis technique (e.g. principal component analysis) could be used to assess quantitatively the coherence of the indicators to be considered together within the same priority activity, (thus indicating to include or to exclude some indicators for the computation of a specific priority activity score) and the coherence of the priority activities to be considered together within the same HCIs (providing, or otherwise, the statistical evidence of the fitness of priority activities for being allocated to a specific dimension or target) (Hair *et al*., 1995).

Moreover, the computation of *Cronbach's α*[28] could provide formal assessment of the consistency of the indicators (or priority activities) in representing the priority activities (or HCIs). Formally, *Cronbach's α* provides a quick measure of the "reliability" of the score, or, in other words, of the relevance of indicators. Cronbach's *α* values provide an ex-post judgment on the quality of the indicators in representing/fitting the overall purpose and dimension of the HCIs. It may also yield a quantitative solution to the potential ambiguity of the sign of the relationship between indicators and the overall HCI, indicating the sign as suggested by the correlation structure of the collected information. Analytically, values of *Cronbach's α* greater than 0.60 may indicate an acceptable level of aggregation of indicators within the same dimension.

4.2. Alternative methodologies for HCI development

As previously illustrated, HCI development necessarily requires steps where subjective choices have to be made. These choices concern the approach for normalizing the indicators as well the approach for weighting and aggregating priority activities for computing the HCIs. In order to assess the extent (if any) of the impact of arbitrary decisions on the final HCI definition, it could be useful to compute HCIs using alternative methodologies, comparing the final outcomes. This step aims to provide further robustness and transparency to the suggested methodology for HCI development, and feasible options for further refinement and optimization of the methodology.

---

[28] Cronbach, L.J. (1951). Coefficient alpha and the internal structure of the tests. Psychometrika. 16, 297-334.

For example, even if this document proposes a qualitative normalization procedure based on quality assessment by NFPs, alternative normalization schemes may be followed. Such schemes could be based on qualitative scales with minimum and maximum values not necessarily equal to those proposed in this document. Alternative quantitative methods for normalization could be followed, for example, ranking the indicators across countries, and then using the country rank, or the categorical scale based on ranking, as dimensionless value. Although this specific approach may not allow the performance of countries across years to be followed with precision, it could provide an empirical basis for assessing the consistency of the subjective assessment provided by the NFP. Moreover, aggregation schemes based on geometric aggregation rather than linear aggregation of priority activities may be tested. Geometric aggregation indeed overcomes the full compensability assumption implied by the weighted arithmetic mean.

Efficient data-driven methods, reflecting the statistical properties of the collected observations, may be further implemented to investigate HCI robustness. In particular, principal component analysis and factor analysis are two widely used approaches for synthesizing the information contained in correlated variables in a smaller set of main dimensions. The weights, in both cases, reflect more or less the statistical quality of the indicator. Weights are indeed the coefficients for the linear orthogonal transformation of the variables for identifying the uncorrelated dimensions.

Mathematical programming provides effective options for overcoming some general limitations of composite index development: subjectivity on the normalization schemes and general disagreement on the definition of weights (Cherchye et al., 2008[29]). Data Envelopment Analysis is based on the estimate of a benchmark or "efficiency frontier" within a space defined by every possible linear combination of priority activity scores. Identification of the benchmark is based on the weight choice obtained as a solution of a maximization problem, where the objective function consists in an ideal point where all the indicators assume the most preferred value. The geometric distance of each country with respect to the efficiency frontier is then used to define the country score, allowing cross-country comparison.

Although widely used in the scientific literature, these data-driven methods are far from perfect and they also require computational effort (time and resources) that may go far beyond the pragmatic nature, clarity and functionality of the HCIs. However, as stated above, they could be taken into consideration as options for future development of the methodology.

5. CONCLUSION

This study proposes a methodology to aggregate ideally multi-faced concepts, corresponding to the 18 priority activities of the Second GPA, into wider and primary dimensions, matching the three mutually supportive targets (PGRFA Conservation, PGRFA Sustainable Use and PGRFA Institutional and Human Capacities) adopted by the Commission at its last Regular Session. It provides guidance and recommendations on how to proceed on the construction of the HCIs. Since some important choices are left to the Commission, the present document highlights the results and consequences of the different strategies with the aim of assisting the decision making process.

The proposed method is driven by two basic principles: the first aims to safeguard the pragmatic nature and functionality of the HCIs. This strictly requires the resulting HCIs to maintain their suitability to address purposes for which they were originally intended: namely, to assess countries' progress towards the three targets, allowing the comparison of performance across countries and regional areas. Thus, HCIs were first conceptualized and then developed as pure instruments in the

[29] Cherchye, L., W. Moesen, N. Rogge, T.V. Puyenbroeck, M. Saisana, A. Saltelli, R. Liska, and S. Tarantola. 2008. "Creating Composite Indicators with DEA and Robustness Analysis: The Case of the Technology Achievement Index." The Journal of the Operational Research Society 59(2):239-251.

hands of stakeholders at all levels to provide a synthetic assessment of the progress towards the three agreed PGRFA targets based on a given set of indicators adopted for monitoring the Second GPA.

The second principle regards the adaptability of the proposed model to measure the complex and multidimensional constructs such as those embracing the 18 priority activities that constitute the Second GPA. Furthermore, the model had to be shaped in such a way as to support the hierarchical/nested structure of the adopted indicators first grouped to priority activities and then in composite indices. The model does not propose changes to the organization of the priority activities or the way they can be measured through the adopted indicators. Indeed, any changes remain a prerogative of the Commission. As regards the data source, it is envisaged that the complexity of the issues under investigation, and the extremely different environmental and social contexts in which countries operate might generate several sources of data heterogeneity, leading to non-comparability of the pure indicators across countries. Thus the proposed methodology has to take into account the potential drawbacks of the 63 indicators, trying to minimize their impact on country scores. Expectation on data availability as well as several sources of data inaccuracy were considered during the development of the proposed methodology for computing HCIs, attempting to ensure consistency, validity and comparability of the three HCIs. The proposed methodology also explicitly takes into account the existence of a wide range of alternative approaches suggested by the international literature as well as the experience gained over several years by researchers in this field.

Analytically, this document emphasizes the relevance of assigning the level of importance to the priority activities of the Second GPA through participatory decision making processes (including the option to assign equal importance to all the priority activities) rather than imposing a choice on the specific way to combine the different dimensions for the development of HCIs. It is proposed that HCIs are expressed through an ordered categorical scale defined on the basis of an expert judgment of the NFP about the country performance and progress over time on each indicator. Calculation of the indicators remains an essential task of the process, since in their assessment the NFPs have to be guided by indicator values. The idea behind this choice is the recognition that only the NFPs may interpret and meaningfully contextualize and codify the quantitative measures coming from the indicators as well as the data collected for compiling them. If, on the one hand, this choice mitigates complications related to the heterogeneity of the values of the indicators referring to particularly different national and environmental contexts, on the other, by adjusting the measured indicators to fit a qualitative common scale, it ensures comparability of the HCIs across countries and time periods.

HCIs can be presented using simple tables of values, including, for each country, individual scores computed for the 18 priority activities of the Second GPA. This will allow the monitoring of national progress on implementation of the Second GPA for each priority activity. Alternatively, following the colour scheme based on traffic-lights developed to measure progress in the implementation of the Global Plan of Action for Animal Genetic Resources, scores can be visualized using eight colours – three shades of red (indicating low levels of implementation), two of yellow (indicating medium levels of implementation) and three of green (indicating high levels of implementation). This type of representation could be used to benchmark country performance over time and to conduct cross-country analysis. Furthermore, the method will assure aggregation at regional and global level.

The main limitation of the proposed model is also its main virtue: its relative generality. As no data sources in line with the adopted indicators were available at the moment of its conceptualization, the model had to be general enough to handle all data availability scenarios. Operative tools for further refinement and optimization of the methodology, once data for the individual indicators become available, were also suggested.

References

Booysen, F. 2002. An Overview and Evaluation of Composite Indices of Development. *Social Indicators Research* 59(2):115-151.

Cherchye, L., W. Moesen, N. Rogge, T.V. Puyenbroeck, M. Saisana, A. Saltelli, R. Liska, and Tarantola, S. 2008. Creating Composite Indicators with DEA and Robustness Analysis: The Case of the Technology Achievement Index. *The Journal of the Operational Research Society* 59(2):239-251

Cronbach, L.J. 1951. Coefficient alpha and the internal structure of the tests. *Psychometrika* 16: 297-334.

Hair Jr, J. F., Anderson, R. E., Tatham, R. L., Black, W. C., 1995. *Multivariate Data Analysis: With Readings*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1995.

Heitjan, D.F., and Basu, S. 1996. Distinguishing “Missing at Random” and “Missing Completely at Random”. *The American Statistician* 50(3):207-213.

Key, V. O. 1940. The lack of a budget theory. *American Political Science Review* 34 (6), 1137-1144.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E. 2005. *Handbook on constructing composite indicators: methodology and user guide* (No. 2005/3), OECD publishing, Paris.

Niemeijer, D., and de Groot, R. S. 2008. A conceptual framework for selecting environmental indicator sets. *Ecological indicators* 8(1): 14-25.

Paruolo, P., Saisana, M., and Saltelli, A. 2013. Ratings and rankings: voodoo or science?. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(3): 609-634.

Saaty, T. L. (1988). What is the analytic hierarchy process?. Springer Berlin Heidelberg.

Saisana, M., and Saltelli, A. 2011. Rankings and Ratings: Instructions for Use. *Hague Journal on the Rule of Law* 3, 247-268.

Saltelli, A. 2007. Composite indicators between analysis and advocacy. *Social Indicators Research* 81(1), 65-77.

Singh, R. K., Murty, H. R., Gupta, S. K., and Dikshit, A. K. 2007. Development of composite sustainability performance index for steel industry. *Ecological Indicators*, 7(3), 565-588.

List of abbreviations used

AHP - Analytic Hierarchy Process

BAL - Budget allocation approach

CGIAR - Consultative Group on International Agricultural Research

EW - Equal weighting

FAO - Food and Agriculture Organization of the United Nation

HCI - Higher-order composite index

MCA - Multi-Criteria Analysis

NFP - National Focal Point

PA - Priority activity

PGRFA - Plant genetic resources for food and agriculture

Reporting Format - Reporting Format for monitoring the implementation of the Second Global Plan of Action for Plant Genetic Resources for Food and Agriculture

Second GPA - Second Global Plan of Action for Plant Genetic Resources for Food and Agriculture

The Commission - The Commission on Genetic Resources for Food and Agriculture