ICT Estimation overview

2022 methodology

September 2023

Data on core ICT indicators are desired for all countries for...

Global comparisons

(1) Track relevant Sustainable Development Goals (SDGs) identified by UN Member States

(2) Monitor differences in connectivity between regions and countries through publications such as ITU's annual Measuring digital development: Facts and figures.

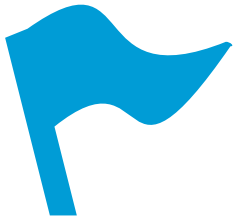
(3) Assure that a sufficient number of countries can be measured through the ITU ICT Development Index

Informed decision-making

Increased awareness of important ICT trends for policy-makers

Directed ICT policies supported by data

Data are incomplete for many countries



146

Member states missing at least one year of official data in recent years

>60%

Share of World population for countries missing data

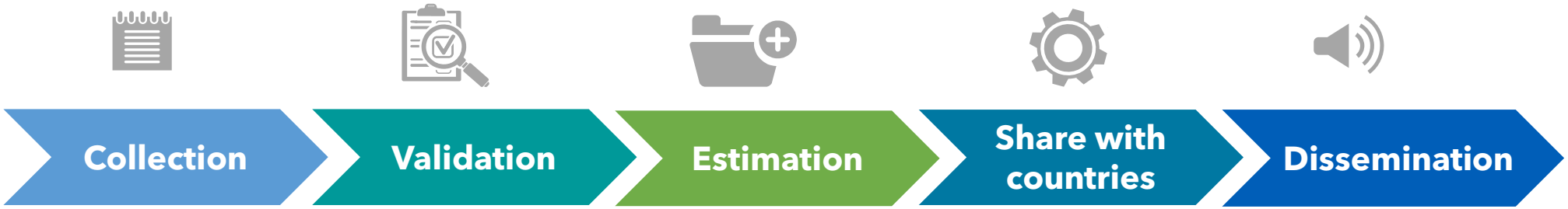
>50%

Share of World GNI for countries missing data

* Internet use data as of July 2023

Far more missing data for household ICT indicators → Cost and complexity of conducting surveys

Estimation timeline



Data request through questionnaires from National Statistics Offices and Government Agencies in charge of telecommunications/ICT (Ministry, Regulator Authority).

ITU validates the consistency of the data received and clarifies with the statistical focal points, (applying statistical techniques, expertise and sectoral knowledge)

Estimation through imputation and modelling where there are data gaps

Estimates shared with countries – where countries disagree with estimates they are not published

Results published via digital channels and reports: ITU DataHub, Digital Development Dashboard, Facts and Figures, Global Connectivity Report.

Short quest. **April** **June** **October** **December** **January** **DataHub update**

2022 estimation cycle

Objective

Estimate and project aggregate and country level data

2019-2022

(1) Share of individuals using the Internet

(2) Share of individuals using the Internet by gender

(3) Share of individuals owning mobile phones

(4) Share of individuals owning mobile phones by gender

(5) Share of individuals using the Internet, by age

(6) Share of individuals using the Internet, by urban/rural

(7) Share of households with Internet access

Overall guidelines

Official data as priority, simple imputation next

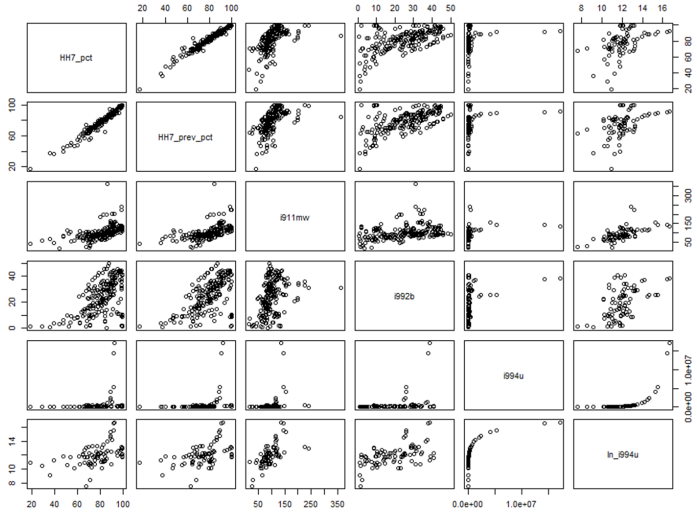
Model remaining gaps in data

Newly available data or other strong evidence needed for major revisions

Estimates for all countries – some unpublished when estimated with low confidence

2022 estimation cycle

Variables considered



ITU indicators

Active mobile-broadband subscriptions per 100 inhabitants

Fixed broadband subscriptions per 100 inhabitants

Prices - Data-only, limited voice and data, basic voice-only baskets - as a percentage of GNI per capita and in PPP

UN Population

Population

Total fertility rate

Median age of population

Life expectancy (overall and gender ratio)

UNESCO

Net enrolment rates (primary and secondary)

Mean years of schooling

Share of schools with Internet access (primary and secondary)

ILO

Labour force participation (15+ and 25+, female participation, gender ratio of participation)

Others

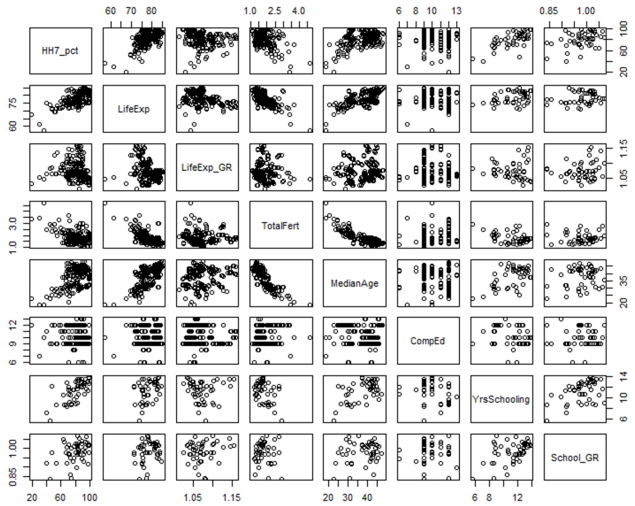
GNI per capita (current US dollars) (UN DESA)

Urban population (percentage) (World Bank)

Cell towers per capita – total, GSM, UMTS, LTE (OpenCellID)

2022 estimation cycle

Other variables considered



Limited availability

International bandwidth per Internet user (bit/s) (ITU)

Price of basic device as a percentage of GNI and in USD (ITU)

Government funding per student (UNESCO)

Literacy rate (overall and youth) (UNESCO)

Government expenditure on education (UNESCO)

Gender parity index (UNESCO)

Special cases

Internet use as base for other estimates

Model development - Official data only for all other estimations

Model implementation - Using country estimates

Internet use by gender, Mobile phone ownership used in a similar way for some estimations

Future consideration

Energy use

Average radiance (controlled by density)

IXPs

Poverty measures

Internet speed (Ookla speedtest)

Dataset comparisons for 2022 estimation cycle

Data availability 2019-2021*

Estimation	Observations	Countries	LDCs
Internet use, overall	**218**	**94**	**5**
By gender	210	89	2
By age	109	53	0
By urban/rural	114	53	2
Mobile phone ownership, overall	**144**	**74**	**5**
By gender	110	64	5

* as of June 2022

Most observations where data are collected in both short questionnaire and long questionnaire – also SDG indicators

2022 estimation cycle

Development of a modelling suite

Regression type

Ordinary least squares (OLS) regression

K-fold cross-validation (k = 10)

Transformation

Linear

Log odds (logit)

Differing datasets

Balance between including useful drivers and maximizing observations

Excluding variables that skew models

Weighted / unweighted

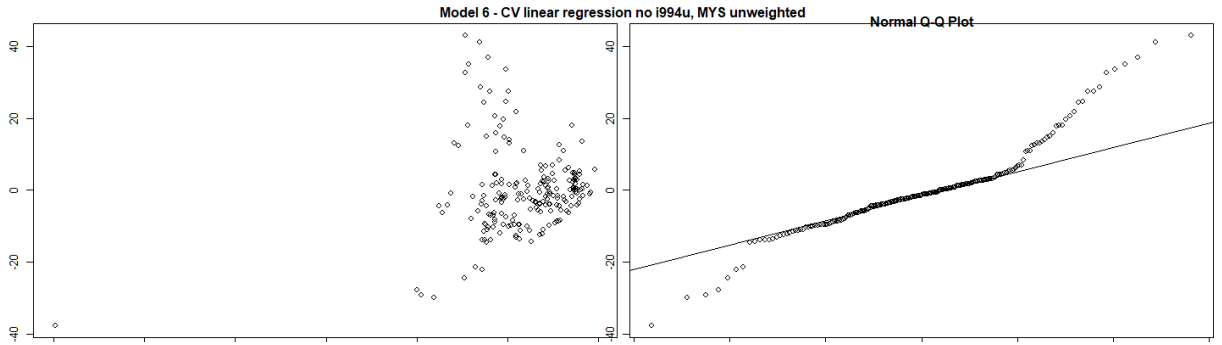
Equal share for each region based on number of countries

Generally resulted in lower estimates and wider gaps compared to unweighted models

Residual analysis for 2022 estimation cycle

Estimation	Estimator	Violations
Internet use, overall	**Overall share**	**22% (4/18)**
By gender	Share of women in Internet users	33% (6/18)
By age	Share of 15-24 users in 15-24 pop	50% (7/14)
By urban/rural	Share of urban users in urban pop	21% (3/14)
Mobile phone ownership, overall	**Overall share**	**23% (5/22)**
By gender	Share of women in mobile phone owners	79% (19/24)

Models excluded more in some estimations than others due to violation of assumptions



Modelling observations for 2022 estimation cycle

Estimator	Share logit transformed	Frequently selected variables
Overall share of Internet users	**24/38**	**(1) GNI pc (2) Population (3) Data year (4) Urban pop share (5) Fixed BB subs**
Share of women in Internet users	13/44	(1) Female pop share (2) Female labour force share and gender ratio (3) Female life expectancy and gender ratio (4) Affordability (5) Mobile BB subs
Share of 15-24 Internet users in 15-24 pop	35/93	(1) Overall Internet use (2) Affordability (3) GNI pc (4) Total fertility (5) Median age
Share of urban Internet users in urban pop	31/103	(1) Overall Internet use (2) Cell towers pc (3) GNI pc (4) Total fertility (5) Median age
Overall share of mobile phone owners	**93/112**	**(1) Overall Internet use (2) Mobile BB subs (3) GNI pc**
Share of women in mobile phone owners	3/124	(1) Share of females in Internet users (2) Affordability (3) Total fertility

Overall use/ownership use logit models more frequently, otherwise linear more frequently chosen

Overall Internet use and GNI pc highly influential in many models – specificities in gender models

Future methodology improvements

New variables

Energy use

Average radiance (controlled by density)

IXPs

Poverty measures

Internet speed (Ookla speedtest)

Methodological

Block cross-validated models to account for lack of independence in time series

Beta regression

Correct for in-scope age differences

Machine learning techniques and/or automation

Estimate earlier years?

More coherence and more information

In coordination with blocked CV models to avoid violation of modelling assumptions

Unwise with fast moving technology changes?

Additional desk research

Some countries have data but do not report

Time consuming – should be done outside of estimation time period

Census cycle offers opportunity to find data for countries not conducting regular surveys

Modeling uncertainty

Future methodology improvements

Input uncertainty

Survey data for dep. variable uncertain

Data for many/all indep. drivers are uncertain

Modelling uncertainty

Uncertainty inherent in statistical modelling

Inputs for models are uncertain and likely correlated

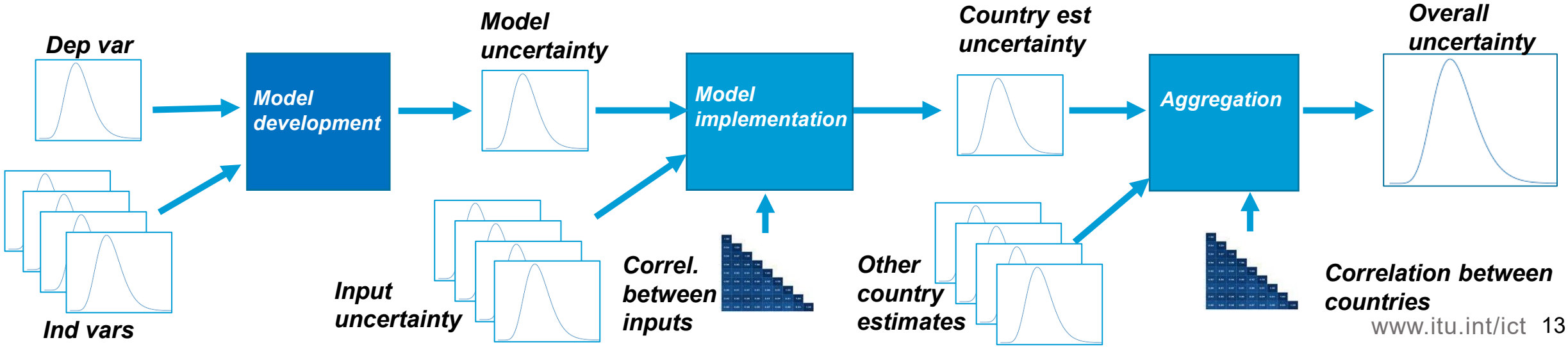
Estimate uncertainty

Resulting country estimates each uncertain and likely correlated

Implementation

Approximated through Monte Carlo simulations

High complexity



Estimates required for a global view of ICT development

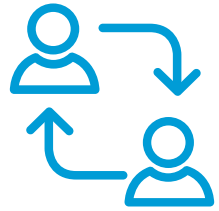
No estimation = many gaps in available data and poor understanding of ICT development in many regions of the world

More data provided = less need for estimation

Limitations

Estimates only as good as data used to develop models – more harmonized data needed especially from low-income countries

Current resources limit the number of estimations that can be performed each year



Progressive improvement

Investigating new methods, ways of work, improved review, and learning lessons from good practices in other organizations

Emphasis can change from year to year depending on member state priorities and feasibility – ICT skills in the future?

Thank you!