



WV GIS Technical Center
Dept. of Geology and Geography
West Virginia University

WV DIVISION OF HIGHWAYS PLAN SCANNING PROJECT

BASIC OPERATIONAL MANUAL

<< **EXTRACT** >>

West Virginia GIS Technical Center
February 9, 2023
Heather Maxey, Annaka Exley

Contents

PROJECT OVERVIEW	1
SEVEN MAJOR PROCEDURAL TASKS:	1
VOCABULARY PRIMER.....	3
FILE PATHS AND PINNING TO YOUR FILE VIEWER'S QUICK ACCESS.....	3
(1) SHEET SCANNING.....	5
SCANNING PROCEDURES FOR 8.5x11 REPORTS	7
SCANNER CLEANING AND MAINTENANCE	8
SCANNING QUALITY ISSUES	8
SCANNER CLEANING.....	8
SCANNER CALIBRATION	8
(2) IMAGE PROCESSING	9
ACTION CREATION PROCEDURES	9
IMAGE PROCESSING PROCEDURES	9
IMAGE SETTINGS AND ENHANCEMENTS.....	11
AUTOMATED ACTIONS	11
(3) PDF BOOK.....	12
PDF RESOURCE NOTES	14
(4) GEOREFERENCING HIGHWAY INDEX MAP	15
ARCGIS PRO SETUP	15
GEOREFERENCE PROCEDURES.....	15
APPENDIX A: Page Number FAQ	20
APPENDIX B: Image Processing Keyboard Shortcuts.....	21

This is a step by step guide intended for internal use by the WV GIS Technical Center to explain the procedures of the DOT Scanning Project. All work progress should be recorded in the appropriate Performance Tracking Database.

PROJECT OVERVIEW

The West Virginia GIS Technical Center (WVGISTC) is working with the West Virginia Department of Transportation (DOT) to scan large format highway plan sheets into digital raster images. For every highway Project Book, a single Adobe PDF file of all the sheets is created along with an index map that is spatially referenced. The WV Division of Highways Plan Scanning Project is fulfilled through a series of steps and processes that are described throughout this documentation. Records are kept to document time spent on each process and are used to calculate estimations of time and cost for future evaluation. Technicians are expected to keep correct records of time spent on each process and record their activity in the DOT_Operator access table and the DOT Performance Tracker Google Sheet.

SEVEN MAJOR PROCEDURAL TASKS:

The scanning project consists of seven major work tasks which are described in more detail below.

(1) Book Preparation and File Naming: Once the Tech Center receives a batch of Project Books from the DOT, book preparation can begin. This process consists of placing each Project Book in the designated location in the Tech Center, keeping them separate from books that have been completed or ones that are file-named and still in progress. Every WV DOT Project Book scanned by the WVGISTC will have an entry in the WV DOT Scanning (*dbo_Scanning*) Database. This database consists of key transportation fields for identifying the highway book, unique scan order and project key identifiers, scan information for tracking project progress and work tasks, index map coordinates, and useful online links for the PDF Book, geographic location, ProjectWise location, etc. After they are placed in their proper location, the technician will select a Project Book and enter description information about the Project Book in the *dbo_Scanning* table. Each DOT Project Book gets an entry in this table during the first intake of the plan. This consists of recording the information on the transmittal sheet (found on top of the map (in the folder/rolled map)) and the Project Book cover sheet. Once all information is documented, the technician will then create and record a Scan Order Number on the transmittal sheet and move the Project Book and transmittal to the location designated for the named documents that are to be scanned. During this, the technician also needs to make sure pages are as neat and organized as possible to speed up the scanning step. This might include repairing pages or replacing binder clips.

(2) Sheet Scanning: The technician will begin this process by selecting a Project Book from the named books area and bringing it to the scanner workstation. The scanner must be set to specific settings before the Project Book can be scanned, which ensures that the image is in the proper format and is named correctly. Additionally, a new folder at the scanner workstation must be created and should be named according to the Scan Order Number on the transmittal sheet. After these are completed, each page of the document is to be wiped off with the cleaning products supplied to protect the scanners from dirt and any other potentially damaging debris. If the technician finds that the Project Book is especially dirty or is in bad condition, document protectors must be used. These are found on the first two shelves of the filing cabinet worktable, which are labelled. Once the technician is sure that the document is clean and the file location and scanner settings are correct, the document can be scanned. After the entire Project Book is scanned, the file must be moved from the local workstation to the IncompleteBooks folder. The physical Project Book should then be noted that it has been scanned and then placed in designated location. *It is important to be sure that all Scanner components are maintained and kept clean.*

(3) Image Processing: This action uses tools created in Photoshop to make image adjustments. The files will be loaded into Photoshop and, with the use of an autocontrast tool, which is detailed below, will be cleaned up, cropped, and rotated, if necessary. After processing, upload the TIFF files into the server folder in the IncompleteBooks folder.

(4) PDF Creation: After image processing has been completed, the edited TIFFs are to be compiled into a PDF document. To do this, the technician uses Adobe Acrobat Pro to reduce file size and recognize text (using Optical Character Recognition). This process is detailed below. After this is completed, the PDF is moved to the PDF folder created on the technicians' local workstation. The document must then be checked and corrected for any rotation or other errors. Once the PDF is correct, rename the document using the Scanning ID Number and put the finished PDF into the server folder in IncompleteBooks.

(5) Georeference Plan Index Map: The technician will use ArcGIS Pro, and Google Maps if needed, to find and georeference each cover sheet of the Project Book. This sheet is typically found on the first page and is contained in the index map, but it can be found in various places throughout the Project Book. The page that has the object to be georeferenced must be copied from the TIFF folder and then pasted into the GeoTIFF folder. Once it is in the proper location, the technician must open the TIFF and ensure the correct projection (Universal Transverse Mercator Zone 17N) is applied. After this is complete, the object's location needs to be found on the map. By using spatial information found throughout the Project Book, the technician can find the location with Google Maps and ArcGIS Pro. Once found, the image is to be georeferenced based off at least four (4) control points set by the technician but must not exceed ten (10) control points. The georeferenced image will then need to be named according to the DOT naming conventions. The GeoTIFF folder is to be moved from the local workstation to the Project Book's file in IncompleteBooks.

(6) Quality Control & File Renaming: Due to the complexity in the several earlier processes, a special emphasis is placed on quality control and assurance to confirm that the technicians are providing the best possible product. The technician will check for accuracy and completion of Database entry, filename, image processing, PDF, and GeoTIFF. Upon completion, the technician will use the file renaming tool to rename the quality assured folder in the IncompleteBooks folder. Once renamed, it will be moved to the ProjectBooksFolders folder. Technicians cannot Quality Control their own work.

(7) Publishing to Web Map: The final process to the WV Division of Highways Plan Scanning Project is uploading the TIFFs, PDF, and GeoTIFFs to the website. First, the PublishGeoTIFF.pyt script will run python to copy and replace GeoTIFFs into a new folder that will be used to add scan rasters to the mosaic. Once moved into a new folder and added to the mosaic, the technician will use the DOTFootprints.pyt tool to create footprints of the georeferenced image. Next, the footprint extents need to be copied into the *dbo_Scanning* table. This is done with the UpdateFootprintExtents tool in ArcMap and will automatically update X and Y Min/Max fields in the *dbo_scanning* table. After this tool is finished, the technician must update the PublicationStatusID column in *dbo_Scanning*, which will publish the projectbook to the images in the ArcGIS Online Map. Once published, the technician checks the website to ensure it works.

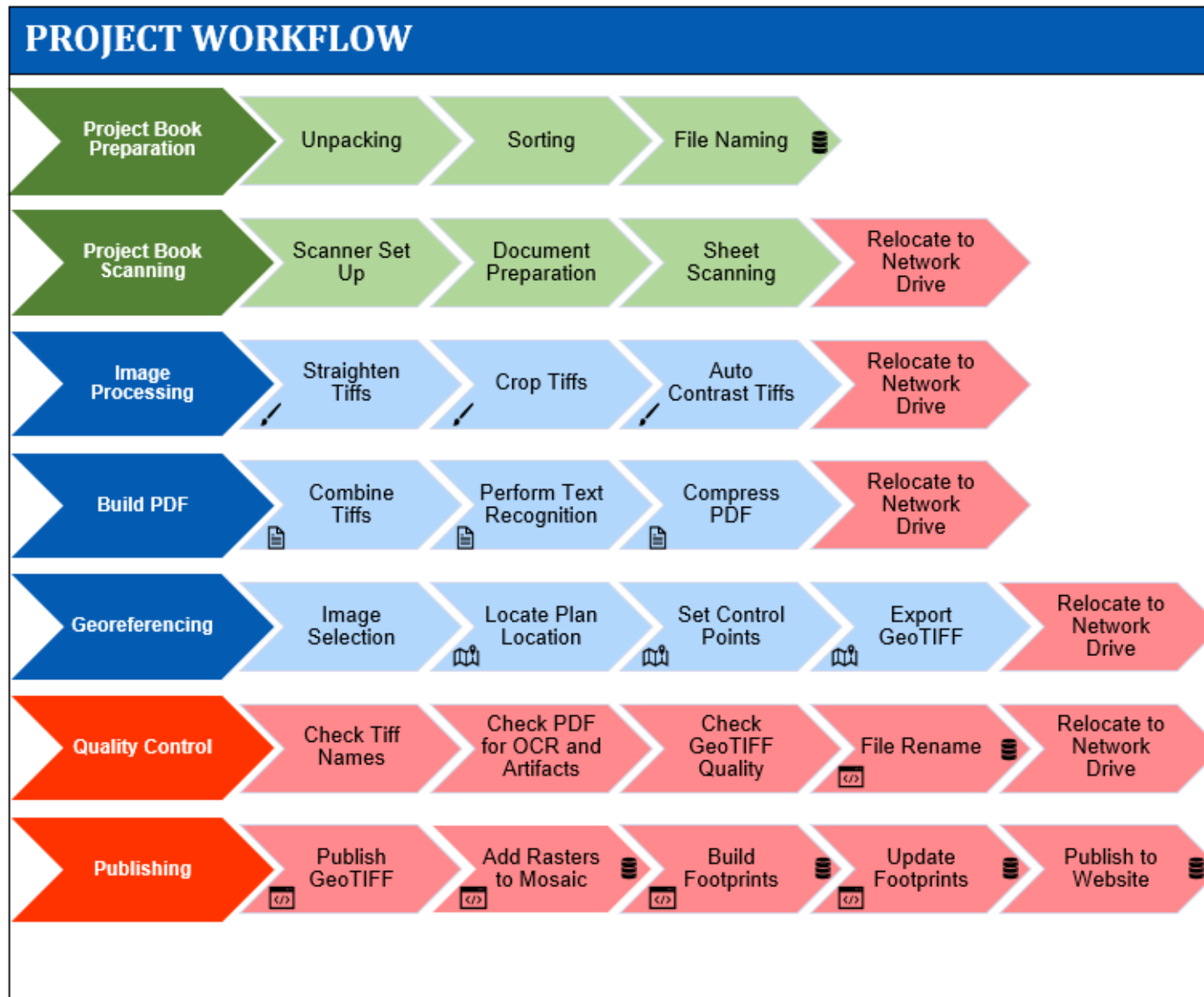
VOCABULARY PRIMER

- **Project Book:** One individual project or plan. This term is interchangeable with 'plan', 'project', or 'book'. These are separated by their transmittal sheets within a folder of projects. Can be from one page to hundreds of pages.
- **Project Book Folder:** The large manila folders of individual Project Books. Folders can have anywhere from one Project Book to many. They need to be kept together and in order of their Scan Order Number. When we return plans to the DOT they are kept in this folder.
- **Transmittal Sheet:** The physical copy of the Project Book information, usually partially filled by the DOT when they are sent to us. Every Project Book gets one. The information on the sheet includes things such as the title of the project and route information which is entered into the DOT Access Database.
- **Scan Order Number:** This is the ID assigned to every individual Project Book and is found on the transmittal sheet. We assign this number based on the order we enter them into the DOT Access database during the naming process.
- **Artifact:** Byproduct produced when there is dirt or other unwanted material on the scanner lens when scanning a page. Usually forms in a streak like line in the direction the page is scanned.
- **TIFF:** The file type the pages of a Project Book are digitally stored as. Stands for Tag Image File Format and is a handy way to store high quality images for editing.
- **Georeferencing:** The process of scaling a plan's map to its real-world location in ArcGIS Pro.
- **GeoTIFF:** The georeferenced index page of a plan. Not all project plans will have the geoinformation to have a GeoTIFF. Usually uses the cover page of a Project Book's map.
- **OCR:** Optical character recognition. Part of the process of building the individual TIFF files into a single PDF book when we run text recognition on the pages to turn the text printed on the plan into a digitally recognized version for document searching.

FILE PATHS AND PINNING TO YOUR FILE VIEWER'S QUICK ACCESS

- **IncompleteBooks:** (<\\gistic-filesrv1\DOTScanning\Scanning\IncompleteBooks>)
 - This is the network drive where all in progress Project Books go.
- **ProjectBookFolders:** (<\\gistic-filesrv1\DOTScanning\Scanning\ProjectBookFolders>)
 - This is the network drive where all complete Project Books go.
- **wvDOT:** (R:\DOT\wvDOT_planROW_scanning)
 - This folder holds the data and management documents for the scanning project.
 - DOT_Operator.accde is for logging your activity.
 - DOT.accdb is the overall database for the project. Do not open or edit anything in this Access sheet without permission.
 - The Documents folder holds the manuals.

To pin the folders to your quick access bar on your file viewer, right click the folder itself and select 'Pin to Quick Access'. It will now be on the left pane of the file viewer.



(1) SHEET SCANNING

Each Project Book will be unbound, and each sheet individually scanned, following all best scanning practices as determined by WVGISTC staff, the scanner's user manual, and the specifications described below.

Each page of a Project Book will be numbered sequentially, scanned in the order it is in the book with the auto naming feature on. This is to aid WVGISTC technicians and staff during the scanning and QC process. Double sided scans will have both sides scanned individually, while trying to remove all visible information from the reverse side if possible. This can be aided by placing black construction paper behind the sheet inside of a protective cover before scanning.

The document-protection scanning sheath is to be used whenever necessary to protect the document and/or scanner. Conditions when this protective sheath is to be used include:

- The document is very dirty and might leave dirt in the scanner
- The document has a waxy coating that may leave residue on the scanner lens
- The edges of the document are torn.
- The document is brittle and could be damaged by the scanner.
- The document is made of a material that is difficult for the sensors in the scanner to properly detect, resulting in some scanning errors/inconveniences
- Any other reason as determined by the scanning technician

Older sheets tend to need the protective sheath more. Newer sheets are often made of clean, sturdy materials that do not require added protection.

SCANNER SETUP

The scanning software used is SmartWorks Pro for both ColorTrac SmartLF SC 42e Xpress Scanners. Turn on the scanner and open SmartWorks Pro scanning software. Ensure the scanner settings are correct.

SCANNING SETTINGS	
Size	22 x 34 Sheets: Full Auto Size 12 x 18 Sheets: Manual Width with Auto Length
Width	22 x 34 Sheets: N/A 12 x 18 Sheets: 480mm
Auto BW points	Off
Color Mode	256 Colors with auto palette (8-bit or indexed color)
Resolution	300 dpi
Quality	1
Auto Naming	Off or Add Numerical Suffix
Auto Overwrite	Type new name if file exists
File Name	S1XXXX-XXXX.tif
File Type	TIFF LZW
Folder	E:\DOT\Imagery\Original Scans\S1XXXX

FOLDER, FILE AND PAGE NAMING

Folder Name: On the scanning computer create a folder named after the Scan Order ID for the Project Book being scanned using the computer pathway listed in the above table.

Scan File Name: The naming convention for individual TIFF digital scans is the Scan Order Number followed by the page or sheet number. Example: Sheet number 25 from Project Book S12345 should have a file name of S12345-0025.tif. Only numbers should be used in the page number, using the auto naming feature of the scanning software.

Sheet/Page Naming: Use the *Add Numerical Suffix* of the scanning software to name the sheets (or pages) in sequential order automatically.

SCANNING PROCEDURES FOR 22x34 AND 12x18 SHEETS

- 1) Select the top Project Book Folder from the named books area (from the most recent shipment) and bring it to your assigned working pile inside the scanning room.**
 - a) If there are many Project Books in the folder, it is recommended to move the Project Books in smaller sections while preserving the order to the top of the working pile assigned to you. This way you can take the Project Books from there to the scanner workstation and back without having to move the whole folder from the scanning workstation if you cannot finish the whole folder in this session.
- 2) Move a Project Book to the scanner workstation.**
- 3) Write the Scan Order Number (e.g. S12345) with a Sharpie marker on the supplied colored paper.**
 - a) The Scan Order Number is found on the Project Book's transmittal sheet.
 - b) There will be one piece of colored paper per individual Project Book.
- 4) Remove clips/tape/staples/sharp objects from sheets and unlatch the binder.**
- 5) Clean the scanner lens.**
 - a) Refer to the cleaning section below for more details.
- 6) Check that scanner settings are correct.**
 - a) Refer to the settings table above. Once you set them for the first time you should not have to change them again but do confirm each time you start scanning.
 - b) On the scanning computer create a folder named using the Scan Order Number from the transmittal sheets. Use the computer path name listed in the table above.
- 7) Scan sheet.**
 - a) Wipe the sheets clean and remove any sharp objects or debris as you go to protect the scanner and improve the image quality of the scan.
 - b) Fix any tears or holes with reinforcement labels or document tape if necessary.
 - i) Use a "D-size" document protector found in the map case if sheets are dirty, excessively waxy or sticky, too torn, brittle or any other reason it can't be scanned like regular sheet.
 - c) Check digital image: Make sure it is straight and free of any streaks or artifacts made by the scanner, folds over the neat lines or other data, no information is cut off, or that there is any other reason you will have to rescan the page for.
- 8) Repeat step 7 until all sheets have been scanned.**
- 9) After the entire Project Book is scanned, the TIFF scans must be transferred from the local workstation to the *IncompleteBooks* folder.**
 - a) The file path can be found in the Plan Overview section.

- b) Mark plan as **Yellow** on the Performance Tracker when the file has been transferred.
- 10) **Replace prong fasteners for the Project Book.**
 - a) Ensure the plan is as securely fastened as possible.
 - i) Use new prong fasteners if the old ones are too damaged or otherwise unacceptable.
 - ii) After scanning, use document hole punch reinforcement labels to secure loose pages with broken prong holes.
- 11) **Place Project Book within large folder that it originally came from.**
 - a) Make sure the Project Books are in order in the folder.
 - b) If you are not finished with a plan, keep it separated from the plans that are finished and the plans that have yet to be scanned in a manner that makes sense to you. Do not leave unfinished folders or Project Books on the center scanning cabinets.
- 12) **When all Project Books in a folder are scanned, write the Scan Order Number range of the Project Book Folder with a Sharpie marker on the colored paper provided and tape it to the large folder.**
 - a) One sheet will be the list of all the projects in that large folder, if there are multiple projects in one folder add the first and last Scan Order Numbers to that list (e.g. S12345 – S12355)
- 13) **Clean the scanner lens when you finish scanning for the day.**
- 14) **Update DOT Performance Tracker Google Sheet:**
 - a) Find relevant Scan Order Number. Mark plan as **Yellow** when file has been transferred if it is not already.
 - i) Date of Action
 - ii) Technician Name
 - iii) Time worked on action (minutes) in the Scanning Time column
 - iv) Quantity completed (total number of sheets)
 - v) Additional Comments
 - b) Complete this process after every completed action.
- 15) **Update DOT_Operator Tracking Database in MS Access:**
 - i) Technician performing action
 - ii) Scan Order Number
 - iii) Action taken
 - iv) Date of Action
 - v) Time worked on action (minutes)
 - vi) Quantity completed (sheets)
 - vii) Additional Comments
 - a) Complete this process after every completed action or before the end of shift.

SCANNING PROCEDURES FOR 8.5x11 REPORTS

Note: Disregard scanning the 8.5 x 11 sheet unless WV DOT includes a note to scan the 8.5 x 11 reports which they consider important. Guidance from Kevin Huffman on 6/16/2017. If you are not sure ask the project lead for guidance and further instructions.

SCANNER CLEANING AND MAINTENANCE

Clean the scanner before and after scanning for the day and recalibrate scanner at least once a month. The lens should be cleaned with a lint-free cloth and specialized lens-cleaning solution. The white rollers on the lid that sit on top of the sensor when the lid is closed also need to be carefully cleaned, and any other dust, hairs, or debris under the scanner lid should be removed.

Occasionally, cleaning the scanner lens and rollers will not improve the quality of the scan, which will continue to have dark streaks, discoloration, or other issues. In these cases, it may help to recalibrate the scanner. The scanner's online manual has detailed instructions on how to perform calibration. If this also does not improve conditions, there may be more serious maintenance issues that will require a visit from a trained technician. It will be important to carefully watch the condition of the scanner over the course of this project so that any issues can be swiftly handled.

SCANNING QUALITY ISSUES

- If images have streaks on them, this means the scanner is dirty. Delete the image, clean the scanner and rescan the sheet.
- If the scanner is unable to catch a particularly crumpled or frayed sheet either rotate the sheet before scanning and rotate it back with Photoshop or place it in a document protector, then scan.

SCANNER CLEANING

- 1) Click in the latches found on the left and right of the base of the scanner to raise the "lid" of the scanner, the scanner must be turned off before cleaning.
- 2) Use air duster to remove particulate material by blowing air across the flat bottom surface and around support and rubberized rollers.
- 3) Remove all residue from all wheels. (Top two tracks of plastic wheels and bottom two tracks of rubber wheels.
- 4) Clean the roller bar. (Distilled water on a magic eraser.)
- 5) Clean the surfaces surrounding the wheels and roller bar to remove any transferred residue.
- 6) Clean the lens and close the lid.
- 7) Wipe down the top of the lid and surrounding surfaces. (Be gentle on the guideline print as it will fade.)

SCANNER CALIBRATION

- 1) Open Smart SC Utilities. On touch screen computer, it can be found at c:\Program Filesx86\ColorTrac\SmartLFSCUtilities\SmartLFSCUtilities.exe.
- 2) Retrieve the calibration sheet from the file cabinet.
- 3) Click the calibration button on screen.
- 4) Screen will prompt you to insert the calibration sheet into the scanner. Insert the sheet exactly as shown.
- 5) Click calibrate and wait for scanner to perform the calibration.
- 6) Close SmartLF Utilities.
- 7) Open SmartWorks Pro.
- 8) Scan the calibration sheet and examine the image to check lines are not skewed. Delete afterwards.

(2) IMAGE PROCESSING

Image enhancements are performed on scanned images using Adobe Photoshop software to improve image quality. Image processing tasks include manual cropping and an automated Auto Contrast function.

Auto Contrast adjusts the overall contrast of an image without affecting its color. Auto Contrast maps the lightest and darkest pixels in the image to white and black, which makes highlights appear lighter and shadows appear darker.

Scans are cropped in Adobe Photoshop to remove excess white space in the margins and to reduce file size. Technicians should crop within an estimated centimeter to the neat line. Technicians **must** be careful to not crop out any information or markings visible on the scan. *When in doubt, do not crop.*

Image enhancement procedures must be performed on the local workstation. Before copying the file over make sure to create the proper subfolder structure first.

Each WVDOT Project Book scanned will have its own project folder for holding the TIFFs, PDFs and GeoTIFFs. The folder name will be the unique Scan Order ID for that Project Book, which is generated by WVGISTC.

ACTION CREATION PROCEDURES

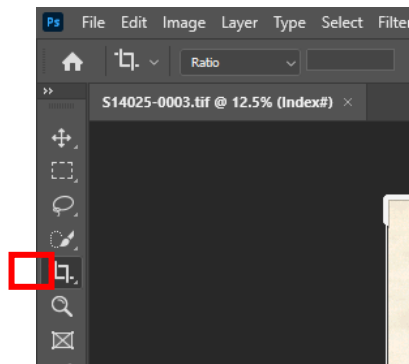
- 1) **Using Adobe Photoshop software, create an action for Auto Contrast and save.**
 - a) Open the Window tab > Actions
 - b) In the Actions menu, click "Create new action" icon at the bottom. (Green box below)
 - c) Name it Auto Contrast and assign a function key (F2).
 - d) Press record.
 - e) Hold down Alt+Shift+Control keys and press L (Alt+Shift+Ctrl+L).
 - f) Then hold down Control+S. (Ctrl+S)
 - g) Click the stop icon in the action menu. (Red box below)
 - h) This will automate the process of using auto contrast and save the file. Use when finished.
 - i) This process will not need to be done again



IMAGE PROCESSING PROCEDURES

- 1) **Open File Explorer.**
 - a) Navigate to the IncompleteBooks folder on \\gistic-filersrv1\DOTProject\Scanning\IncompleteBooks
 - b) Open the Project Book to be image processed.
 - c) Create three new folders on your desktop. You will not have to do this more than once. Name them as follows:
 - (1) GeoTIFF

- (2) PDF
- (3) TIFF
- ii) For every Project Book you do, copy these folders from your desktop and paste them into the Project Book folder in the IncompleteBooks folder on the network drive.
- iii) Move the original scans into the TIFF folder.
- 2) **Copy the Project Book files from IncompleteBooks on to your desktop.**
- 3) **Open the tif files in the TIFF folder in Photoshop by selecting them in the folder on your desktop and dragging them to the program.**
- 4) **Making sure you are on the Crop tool, highlighted below.**
 - a) Shortcut 'C' - see Appendix B for more shortcuts.
 - b) Align the page by holding Ctrl or clicking the "ruler" icon and dragging your mouse from one corner to another on the straightest line possible.



- c) Ensure the text is as level and straight as possible.
- d) If there is no neat line, use your best judgement to find the next best reference.
- 5) **Crop the image to around a centimeter (estimate it) away from the neat lines by using the crop tool while making sure to include any text, diagrams, or additional information that is outside the neat lines while minimizing whitespace and the scanner background.**
 - a) If needed, you can use the zoom tool (Z) to get a more precise crop by zooming in closer to the edge. Ctrl + 0 will refit the image to the screen.
 - b) If there is no neat line, keep the cropping about a centimeter away from any information on the page while leaving no white space.
 - c) Keep each side of the page as consistent as possible distance wise, excluding any information outside the neat lines.
- 6) **Use the Autocontrast and Save tool by hitting the function key (F2) you assigned it to.**
- 7) **Go to the next page by hitting ctrl+tab or clicking the next file at the top of the work area.**
- 8) **Repeat steps 4-7 until finished with all the TIFF files.**
 - a) As you go, make sure there are no artifacts, folds over the neat lines or other data, no information is cut off, or that there is any other reason you will have to rescan the page for.
- 9) **Upload processed images to IncompleteBooks from the folder on your desktop.**
- 10) **Update DOT Performance Tracker Google Sheet:**
 - i) Find relevant Scan Order Number.
 - ii) Date of Action
 - iii) Technician Name
 - iv) Time worked on action(minutes) in the Processing Time column
 - v) Additional Comments
- b) Complete this process after every completed action.
- 11) **Update DOT_Operator Tracking Database in MS Access:**

- i) Technician performing action
 - ii) Scan Order Number
 - iii) Action taken
 - iv) Date of Action
 - v) Time worked on Action (minutes)
 - vi) Quantity completed (sheets)
 - vii) Additional Comments
- a) Complete this process after every completed action or before the end of shift.

IMAGE SETTINGS AND ENHANCEMENTS

Adjusting the image levels is also useful for removing background noise in the whitespace of a scan, and for darkening lines slightly. With proper use of keyboard shortcuts, actions, and practice, this step can be performed very quickly and may result in much higher image quality. This process can be automated for multiple images that require identical, small adjustments. Any TIFFs that might be georeferenced later need to be in Indexed color mode and not grayscale mode. They should be in the mode by default if the scanner settings were properly set up.

IMAGE QUALITY ENHANCEMENT	SETTINGS OR ENHANCEMENT TOOLS
Color Mode	Indexed Color Mode (256 colors)
Auto Image Edits	Auto Contrast or Auto Levels
Cropping	Manual

COLOR MODE

The color mode or image mode determines how colors combine based on the number of channels in a color model. To reduce file size, the color mode of the image should be Index Color mode (256 colors). Indexed Color mode produces 8-bit image files with up to 256 colors and is enough colors to match the original document.

AUTOMATED ACTIONS

An action is a series of tasks that you play back on a single file or a batch of files—menu commands, panel options, tool actions, and so on. For example, you can create an action that changes the size of an image, applies an effect to the image, and then saves the file in the desired format.

Actions can include steps that let you perform tasks that cannot be recorded (for example, using a painting tool). Actions can also include modal controls that let you enter values in a dialog box while playing an action.

Photoshop comes with predefined actions installed that help you perform common tasks. You can use these actions as is, customize them to meet your needs, or create new actions. Actions are stored in sets to help you organize them. See the Adobe Photoshop online help or YouTube videos for more information on how to create actions.

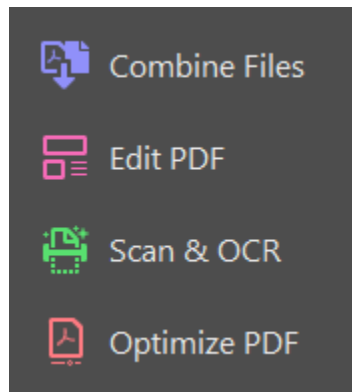
(3) PDF BOOK

After image processing has been completed, the edited TIFFs are compiled into a PDF document. To do this, an automated script created in Adobe Acrobat Pro is executed to reduce file size and create OCR.

For each WVDOT Project Book a single PDF document is created that will hold every scan in that Project Book and follows the correct page order.

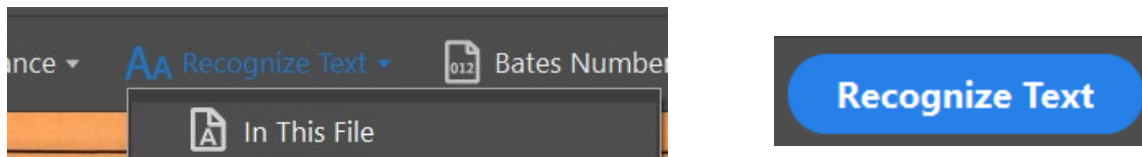
PDF CREATION PROCEDURES

- 1) **Open Adobe Acrobat Pro.**
- 2) **Change your display theme to dark by going to the view tab on the top left menu ribbon > Display Theme > Dark Grey.**
 - a) You only must do this once. It is important to change this setting to see errors later in the process.
- 3) **Open the Tools tab on the top right of the window.**
 - a) On the right side of the screen remove every tool except for Combine Files, Edit PDF, Scan & OCR, and Optimize PDF. Your tool bar will look like this:



- 4) **Open the folder with the processed images from the copy of the plan on your desktop.**
 - a) *Do NOT work out of IncompleteBooks.*
- 5) **In Adobe Acrobat, click on Tools > Combine files.**
 - a) Generally, Acrobat can handle ~50 pages of the 22x34 sheets and ~100 of the 12x18 sheets. This number can vary.
 - b) If it cannot handle the number of pages, split the project into however many PDFs as needed. (e.g. 400 small sheets should have four 100-page PDFs.)
- 6) **Highlight all the processed files you want to combine into a PDF in the TIFF folder.**
 - a) Drag and drop them into Adobe Acrobat with the Combine tool open.
- 7) **Make sure all files are in the correct order, then click Combine.**
 - a) The PDF will then be combined creating a tab entitled "Binder1.pdf".

8) In the OCR tools, run recognize text tool > In this File > Recognize Text (Blue button.)



9) Save and name the PDF to match the number of the Project Book (e.g. S12345.)

10) After that is finished, open the File Optimization tool and run Reduce File Size or Compress PDF depending on your software version.

11) Save and check all the pages for errors.

- a) Make sure that all the pages are the right orientation.
 - i) Open the thumbnail view on the left side and use the rotate arrows to correct the page's orientation.
- b) Make sure there are no OCR errors generated from the text recognition process.
 - i) The page will look like it has been rotated, sometimes barely noticeable and sometimes extremely obvious. These pages will either need cropped to remove the generated white space or replaced if cropping is not possible.
 - ii) Open the thumbnail view on the left side and right click the page that has the error. Replace the page with the processed original page from the TIFF folder. You will have to change the file type in the file viewer to All or TIFFs to see the files. Press ok when prompted.
- c) Make sure there are no artifacts, folds over the neat lines or other data, no information is cut off, or any other reason you will have to rescan and reprocess the page for.
 - i) In this case, find the page, rescan it, replacing the page in the OriginalScan Project Book folder.
 - ii) Upload that individual file to the IncompleteBooks folder and replace the former scan.
 - iii) Process the new page in Photoshop as usual.
 - iv) Replace the old page in the PDF with the new page you rescanned.

12) If there are multiple PDFs from too many pages, repeat steps 5-11 till all pages have been built into a PDF.

- a) Combine them into one singular PDF using the Combine tool, making sure they are in the right order. There should only be one PDF in the incomplete books folder.
- b) Save and name the PDF to match the number of the Project Book (e.g. S12345.)
- c) Re-run the Reduce File Size or Compress PDF tool.

13) Upload finished PDF to IncompleteBooks from your desktop.

14) Update DOT Performance Tracker Google Sheet.

- a) Find relevant Scan Order Number.
 - i) Date of Action
 - ii) Technician Name
 - iii) Time worked on action(minutes) in the PDF Time column.
 - iv) Additional Comments
- b) Complete this process after every completed action.

15) Update DOT_Operator Tracking Database in MS Access:

- i) Technician performing action
- ii) Scan Order Number
- iii) Action taken

- iv) Date of Action
- v) Time worked on Action (minutes)
- vi) Quantity completed (sheets)
- vii) Additional Comments
- a) Complete this process after every completed action or before the end of shift.

PDF RESOURCE NOTES

OCR Sheet Rotation: The automated Recognize Text process involves a step that automatically straightens the image by rotating it so that most of the text lies perfectly horizontally. This occasionally results in errors, such as the image being rotated to match non-horizontal text so that the scan now appears crooked, or the image being rotated a full 90 degrees to match text that appears vertically. Because of this, each page of the PDF needs to be inspected after the Recognize Text tool is run, and any errors corrected. For slightly crooked images, the correction will require removing the erroneous page and inserting it into the correct place within the PDF WITHOUT running the Recognize Text tool on that page (since that will result in it being crooked again).

(4) GEOREFERENCING HIGHWAY INDEX MAP

Technician will use Photoshop, Google Maps, and ArcGIS to find and geo-reference each Project Book index map to **UTM Zone 17N projection, NAD83 Datum**. The target object to be georeferenced is typically found on the first page and is contained in the index map, but it can be found in various places throughout the Project Book. The page that has the object to be georeferenced must be copied from the TIFF folder and then pasted into the GeoTIFF folder.

ARCGIS PRO SETUP

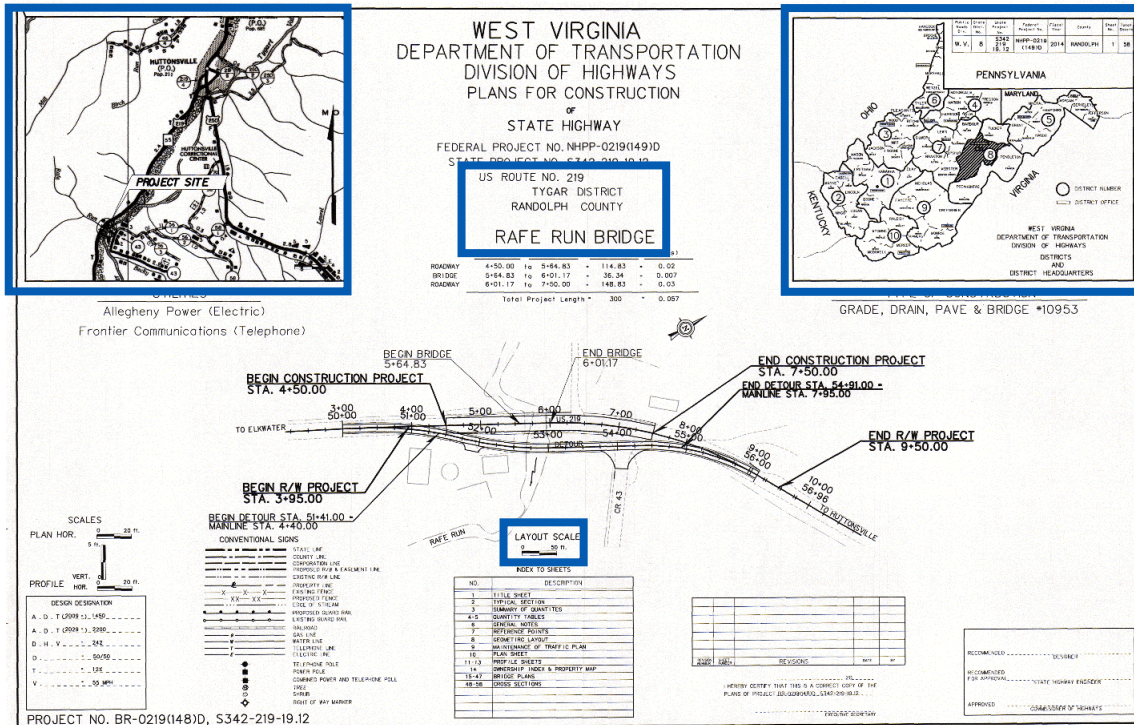
- 1) Open ArcGIS Pro and log into your account.**
 - a) Make sure the program is up to date. You might need admin permissions to do this, so ask the project lead for help.
- 2) On the start screen, create a new map by clicking on the first icon labeled “Map”.**
 - a) When prompted, name the project “DOT_GR”. This will now open the ArcGIS interface.
- 3) Once the interface is loaded, click the arrow underneath “Connections” in the Insert tab, hover over the “Server” option, and click “New ArcGIS Server”**
- 4) Paste this link in beside the “https://”:**
 - a) *services.wvgis.wvu.edu/ArcGIS/services*
 - b) Press OK.
- 5) On the right of the screen should be the catalog pane.**
 - a) If not, click View > Catalog Pane.
 - b) There, click the server folder option, and then the ArcGIS server you added. Open that and the Imagery_BaseMaps_EarthCover folder.
 - c) Drag and drop the wv_imagery_WVGISTC_leaf_off_mosaic into the contents pane on the left.
- 6) Right click “Map” under drawing order on the contents pane on the right of your screen and click properties.**
 - a) Click Coordinate systems in the map properties window. Navigate to Projected Coordinate Systems > UTM > NAD1983 > NAD1983 UTM Zone 17N. Click on it and press ok.
- 7) Once out of the properties window, change your base map to the “Streets” option on the Map tab of the tool ribbon using the Base Map button.**
- 8) Save the project and find the folder in C:\Users\YOUR USERNAME\Documents\ArcGIS.**
 - a) Right click the “DOT_GR” folder and pin it to your quick access bar as instructed in the beginning of the manual.
- 9) You will only need to do this set up once.**

GEOREFERENCE PROCEDURES

- 1) Select the best single TIFF file from the TIFF folder on your desktop.**
 - a) This is preferably the cover page, but not always.
 - b) Open the file in photo viewer.
- 2) Copy the selected file into the GeoTIFF folder on your desktop.**
 - a) Exception to the cover page include:
 - i) Maps that are larger than 1 inch to 3 miles scale or it covers more than 3 counties.

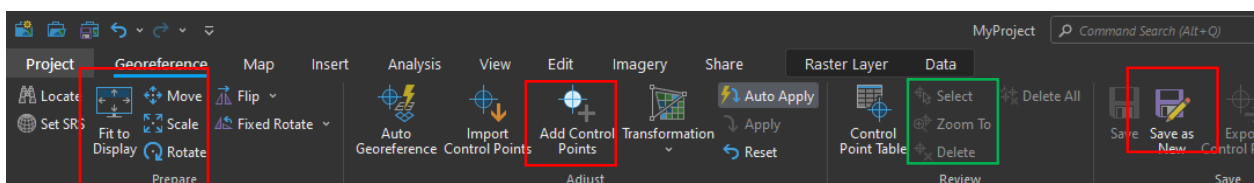
- (1) If between 2-3 miles or 1 inch to 2000ft and 3000ft, crop the page down to just the map on the cover page.
 - ii) Maps without cover pages. Find the next best page that shows as much of the plan as possible and use that.
 - iii) Maps without any geo-information of a reasonable scale. The last resort is to crop out the inset map and georeference that.
 - iv) Maps without any geo-information at all. Comment on the Performance Tracker that the plan has no geo-info. Move to step 15.
- 3) **Drag the TIFF into the Table of Contents from the GeoTIFF folder on your desktop.**
- a) A prompt will appear asking if you would like to create pyramids and statistic information. Check the boxes to have it automatically perform this. Click yes. You will only have to do this once.
 - b) Make sure to place it above the imagery layer so you can see the plan.
- 4) **In the Raster Layer tab and change the Transparency to ~30%.**
- 6) **Use the image you opened earlier to look for spatial references to match to the image in Arc such as roads, rivers, crossroads, etc.**
- i) You can use Google maps to aid you in looking for the area if needed.
 - ii) Sometimes inserting the map and orienting it properly can help you find it. Refer to step 7 on how to do this.
- a) First, find the county the plan is in by referencing the top right corner map. Zoom to this county in ArcGIS.
 - b) Next, use the inset map to find the more precise location in the county where the project is. It will often be depicted by a line, arrow, or circle.
 - c) Once you have the general location, refer to project plan itself.
 - d) For sheets depicting little to no spatial information, use these clues to aid in georeferencing:
 - i) A North Arrow will aid in orienting the page correctly (Top left corner of example below.) Rarely the north arrow will be oriented incorrectly but normally it is correct.
 - ii) Project Title often has more information corresponding to the location of the project, i.e. "Rafe Run Bridge."
 - iii) The scale bar can help you determine how small the plan is.
 - iv) The DOT Access database "dbo_Scanning" can also supply more information concerning any routes/sub routes and other information. Ask the project lead for this information or consult the transmittal sheet attached to the physical Project Book.
 - v) Another tip for reading DOT book plans. Roads that are dashes are generally the old roads and the solid line with the circles and dashes are the proposed plan and should be georeferenced rather than the dashes, unless it does not match.
 - e) If the cover sheet does not supply enough spatial information, try to find a sheet in the project that does, but it should be a last resort.

- i) If there are no sheets that can be used, leave an empty GeoTIFF folder in the Project Book folder. Leave a comment saying “No Geoinfo” in the Performance Tracking Sheet.
- 7) Once you have found the area to which you plan to reference the image click **Georeference**



from the Imagery tab and click **Fit to Display**.

- a) Use the Rotate tool on the image to orient it using the north arrow and use the Move and Scale tools to help you get it into the general position. Do not go for perfect accuracy as that can make it more difficult to add control points.
- 8) Once set in place, click the “**Add Control Points**” button and add control points as necessary.
- a) Always assign control points using the imagery and not the base map.
- b) When you first click to add a control point, click on the section of the map itself and *then* on the corresponding point on the imagery.
- c) Add *at least* four (4) control points and make sure the main portion of the project (Usually outlined in sections of boxes) is what is being georeferenced (e.g. if the plan is an Interstate, match the Interstate road, and not the side roads.) *Do not exceed* ten (10) control points.
- d) Sometimes a point will not work as you wanted or will need removed.
- i) To do this, select the point by clicking the select tool in the review section of the Georeference tab. Once the point is selected, click on the Delete tool. These tools are highlighted in green.
- ii) If you only have one other point placed, this will transform the image to make it ridiculously small. You can either reset the image and start over or try to carry on if you think you can match the point well enough. It will automatically transform the image for you.



9) General Tips:

- The north arrow is not always oriented perfectly or correctly even, if you are having difficulty finding a location, try orienting the map in different directions.
- Check that you are trying to georeference the correct line on the map. Often the plan will be misleading and show the old part of the project in a more precise manner as the part you are most likely georeferencing had not been built yet. Generally, they will represent the new part with a bold line with dashes and circles throughout. Sometimes they will label it but not always.
- The scale bar can give you a hint as to how small or large the plan needs to be. Using the scale of the map on the bottom of ArcGIS Pro you can tell how zoomed in you should be before you fit the map to the display.
- Do not try to make every map as perfect as possible. Some maps are not scaled correctly. Sometimes it is the whole map, and sometimes it is part of the map. There are also times where the plan was not built as planned or perhaps not at all. Obviously, we want the best match we can get, but redoing the map over and over is a waste of your time.
- If you cannot find the location, ask for a second opinion. This is a skill that you develop by doing and will take time to get the hang of.

10) Once done, click “Save as New” in the Georeference tab

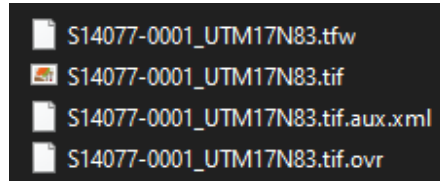
- Set the Output Raster Dataset to your project folder that should be pinned to your quick access bar. It should default to this.
- Make sure the coordinate system is set to NAD_1983_UTM_Zone_17N.
- Leave Geographic Transformations, Clipping Geometry, and Cell Size as it is.
- Change the raster size columns and rows to 6000. The cell size will change when columns and rows changes. That is fine.
- Leave Pixel Type alone.
- Change NoData to 0.
- Leave everything else as it is.
- Click Export
- This image has what it should look like:

The screenshot shows the 'Export Raster' dialog box in ArcGIS Pro. The following settings are highlighted:

- Output Raster Dataset:** C:\Users\{USERNAME}\Desktop\GeoTIFF\ST (highlighted with a red box)
- Coordinate System:** NAD_1983_UTM_Zone_17N (highlighted with a red box)
- Geographic Transformations:** None (highlighted with a red box)
- Clipping Geometry:** Default (highlighted with a red box)
- Cell Size:** X and Y (highlighted with a red box)
- Raster Size:** Columns: 6000, Rows: 6000 (highlighted with a blue box)
- Pixel Type:** 8 Bit Unsigned (highlighted with a red box)
- NoData value:** 0 (highlighted with a blue box)
- Renderer Settings:** Force RGB, Use Colormap, Use Renderer (all unchecked)
- Output Format:** GRID (highlighted with a red box)
- Compression Type:** None (highlighted with a red box)
- Compression Quality:** (highlighted with a red box)

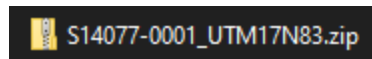
11) When exported, you will find four files in the project folder pinned to your quick access bar on your file viewer.

- a) Rename them all to S1XXXX-[FILENUMBER]_UTM17N83 (e.g. S10899-0001_UTM17N83).
 - i) The bottom two files will select the extra file extensions, make sure you keep all of them and only rename the file itself.
- b) **It is a hyphen between S1XXXX and [FILENUMBER] and an underscore between [FILENUMBER] and UTM17N83. The [FILENUMBER] is FOUR digits.**



12) Select all four files, right click on the .tif file with them all selected and click “Send to Compressed (Zipped) Folder”.

- a) Make sure the zipped folder is named the same as files and the file type is ONLY “.zip”. (e.g. S02939-0028_UTM17N83.zip)



13) Put the zipped file into the Project Book GeoTIFF folder on your desktop.

- a) Delete the copy of the .tif and the two files generated in the GeoTIFF folder by ArcGIS Pro when you loaded the image into the map.
- b) Delete the four tiff files and the .zip file from the DOT_GR project folder.

14) Upload finished GeoTIFF to IncompleteBooks from your desktop.

15) Update DOT Performance Tracker Google Sheet.

- a) Find relevant Scan Order Number. Mark the number as **green**.
 - i) Date of Action
 - ii) Technician Name
 - iii) Time worked on action(minutes) in the Georeferencing Time column.
 - iv) Additional Comments
 - (1) If the plan was particularly difficult to georeference, explaining why the specific location was chosen.
 - (2) Leave a note if the page could not be georeferenced and mark time as N/A.
- b) Complete this process after every completed action.

16) Update DOT_Operator Tracking Database in MS Access:

- a) If there was no Geoinfo, do not log anything in the Access database.
 - i) Technician performing action
 - ii) Scan Order Number
 - iii) Action taken
 - iv) Date of Action
 - v) Time worked on Action (minutes)
 - vi) Quantity completed (sheets, always one sheet for georeferencing.)
 - vii) Additional Comments
- b) Complete this process after every completed action or before the end of shift.

APPENDIX A: Page Number FAQ

Page Numbering FAQ

The following Q&A's are examples of page numbering questions/inconsistencies that have arisen during the course of the scanning project. New questions are added to the document as problems are discovered, and answers are added as they are received.

Q: Pages are arranged out of order.

A: Rearrange the pages by the official page number as seen on the sheet. If there are pages from different projects within the same book which are numbered by a different system/sequence, retain the order in which those pages were physically found when creating the PDF document, as they were likely placed in that order intentionally by WV DOT.

Q: Two pages with different content, but same page number.

A: They might be from different projects. If so, retain the order in which they were found, keeping the same file name and allowing the scanning software to auto number the pages.

There is also a chance the pages were miss-numbered. Example: 17, 18, 20, 20... In a case like that, correct the page numbers in the file name (in the example, you would change the first 20 into 19), retaining the order in which the pages were found.

Q: Some pages are marked as deleted with a large X crossing the entire page and text that reads "This sheet deleted, see page 81A" or the like. In this example, there was a page 81, then another page 81 that was marked as deleted and referred to page 81A.

A: The deleted page does not need to be scanned (according to WVDOT), though it may be wise to scan it anyway. Make a note of it in the database either way.

Q: Some sheets have no visible page number.

A: Give them one for now, according to the sequence in which they were discovered. Keep them in the order they were discovered when creating the PDF document. Make a note in the comments field of the database describing how they were numbered.

Q: One project book can contain sheets from multiple projects. EXAMPLE: Project 32, starting on the second page it switches to a different project in a different county for 37 pages before switching back to the original project listed on the cover sheet, beginning with page #2. Another EXAMPLE: Project 17 has two pages appended to the end from unrelated projects. Another EXAMPLE: Project 33 , page 52 (there are two of them) is a seemingly random page from some unrelated project. How should we handle examples like these? Separate them? Keep them in the same folder and number them as they appear?

A: Those pages were usually put in there by WV DOT staff for a reason, so be sure they are arranged in the PDF as they were discovered. As for naming the TIFFs, they are to be auto named based on the order they are in and still remain inside the project folder in which they were found. Write a descriptive note in the comments field of the database.

Q: Pages repeat, but with some revisions.

A: The original un-revised page does not need to be scanned (according to WVDOT), and the revised page should get the original page number. Scan the un-revised page anyway and note it in the comments field of the database.

APPENDIX B: Image Processing Keyboard Shortcuts

The following shortcuts can be used in Adobe Photoshop for Image Processing:

Shortcut	Action
Ctrl + C	Copy
Ctrl + V	Paste
Ctrl + X	Cut
Ctrl + S	Save
Ctrl + Zero	Zooms to Page Extent
Z	Switches to Zoom Mode
C	Switches to Crop Mode
Ctrl + Tab	Go to Next Image
Ctrl + Alt + Shift + L	Auto Contrast
Ctrl + Alt + Z	Undo Droplet Action
HOLD Ctrl	(While in Crop Mode) Use Straighten Tool
Ctrl + Z	Undo
Ctrl + Plus	Zoom in
Ctrl + Minus	Zoom out
Ctrl + L	Opens Levels Menu

Additional shortcuts may be created using the Keyboard Shortcuts menu in Photoshop (Edit<Keyboard Shortcuts... or Alt + Shift + Ctrl + K)